5-12-2012

Synthesis of fluorogenic substrates for the enzymatic characterization of RV0045c from M. tuberculosis

Kelly Jo McKenna
Butler University

BUTLER UNIVERSITY HONORS PROGRAM

Honors Thesis Certification

Please type all information in this section:

Applicant: Kelly Jo McKenna
(Name as it is to appear on diploma)

Thesis title: Synthesis of fluorogenic substrates for the enzymatic characterization of RV0045c from *M. tuberculosis*

Intended date of commencement: May 12, 2012

Read, approved, and signed by:

Thesis adviser(s): R. Jeremy Johnson — Date: 5/2/12

Reader(s): Geoffrey C. Thorpe — Date: 5/2/2012

Date

Date

Certified by: Judith Harper Morrel, Director, Honors Program — Date: 12 July 2012

For Honors Program use:

Level of Honors conferred: University: Magna Cum Laude

Departmental: University Honors Program

Synthesis of fluorogenic substrates for the enzymatic characterization of RV0045c from *M. Tuberculosis*

A Thesis

Presented to the Department of Chemistry

College of Liberal Arts and Sciences

And

The Honors Program

of

Butler University

In Partial Fulfillment of Requirements for Graduation Honors

Kelly Jo McKenna

May 2012

TABLE OF CONTENTS

TABLE OF FIGURES, EQUATIONS, AND TABLES

Background and Significance

Results and Discussion

Materials and Methods

Abstract

Mycobacterium tuberculosis is the pathogenic bacterial agent commonly responsible for tuberculosis, or TB. Although treatment exists for the active form of tuberculosis, no method has been developed for eliminating *M. tuberculosis* in its dormant state. One hypothesized method for the elimination of dormant TB is to develop an inhibitor specific for *M. tuberculosis* esterases and lipases, as these esterases and lipases are essential to the survival of dormant TB infection. In this research, the substrate specificity of the Rv0045c esterase from *M. tuberculosis* was studied due to the essential role of Rv0045c in TB metabolism and its dissimilarity to other esterases. Two fluorogenic substrates, fluorescein di(ethylbutyloxymethyl ether) and fluorescein di(phenylbutyloxymethyl ether), were designed to match the binding pocket structure of Rv0045c and to test the substrate specificity of the Rv0045c esterase. Both fluorogenic substrates were synthesized *via* a three step synthetic process with reasonable yields (1.63% and 5.6%) and purified using column chromatography. Correct purification of one final product was confirmed using liquid chromatography mass spectrometry (LCMS) and HNMR, giving the expected masses and NMR spectrum. This substrate was then tested for its stability to hydrolyze in water and the kinetics for activation by Rv0045c. Michaelis-Menten kinetic values were determined for the interaction between the substrate and both Rv0045c and pig liver esterase (PLE), a common esterase. The specificity constants for Rv0045c and PLE were 2.95 M^{-1} s^{-1} and 20.99 M^{-1} s^{-1}, respectively; these results indicate that the synthesized substrate shows no specificity to Rv0045c and that it would be inappropriate to use for the development of a prodrug against TB.

Background and Significance

Tuberculosis and Mycobacterium tuberculosis

Mycobacterium tuberculosis is the pathogenic bacterial agent commonly responsible for causing tuberculosis (TB). TB is highly contagious and is spread through droplets either from the throat or lung of an actively infected individual. According to the World Health Organization (WHO), there were 9.4 million new cases of TB and 1.7 million deaths from the infection in 2010 (WHO 2010). Although a 6-month antibacterial treatment for TB exists, there are two issues that prevent its elimination: TB primarily affects those in developing countries with limited access to treatment, and the antibacterial treatment can only be used once the immune system weakens enough to allow symptoms to show (Cotes *et al.,* 2008). These limitations demand that other methods of treatment be pursued.

Interestingly, many patients infected with TB are unaware that it is in their system because *Mycobacterium tuberculosis* has the capability of lying dormant within the body for decades without showing any symptoms in the patient (Cotes *et al.,* 2008). This dormancy is triggered by the presence of the patient's immune system, and it allows the pathogen to essentially 'hibernate' until the immune system is weakened and the opportunity presents itself to attack. Currently, the antibacterial treatment for TB patients is only effective in killing *Mycobacterium tuberculosis* when it is in its active state and may not eliminate all bacterium if there are some that lie dormant during treatment; a new method of eradicating the pathogen while it lies in its dormant phase is desired in order to prevent an active TB infection (Cotes *et al.,* 2008).

One of the ways in which *M. tuberculosis* is able to lie dormant for long periods of time is through lipid metabolism. Analysis of its genome shows that about 250 genes are used to code enzymes for this function, whereas *Escherichia coli* only contains about 50 genes toward lipid metabolism (Cole *et al.,* 1998). These genes are proposed to code for lipid metabolism enzymes that are used both intracellularly to release stored energy from lipids and fatty acids and extracellularly to breakdown the membranes on host cells (Cotes *et al.,* 2008). This intracellular and extracellular enzyme activity is necessary for *M. tuberculosis* to survive in its dormant state, and thus survive until the host immune system is sufficiently weakened. Inhibition of enzymes critical to lipid metabolism is a potential method being researched in this thesis as a cure for TB. Ultimately, this thesis will examine the substrate specificity between a TB esterase and two designed substrates that could potentially assist in esterase inhibition.

Enzymes

Proteins that are used to increase the rate of a chemical reaction are called enzymes. Enzymes act as catalysts, meaning that they speed up reactions without being altered by the reaction; this unaltered structure allows for an enzyme to be reused again and again during its lifespan (Nelson & Cox, 2008). Enzymes contain pockets called active sites. These active sites bind one or more molecules, called substrates, which are involved in the chemical reaction being accelerated. The enzyme uses substrate positioning or conformation changes in order to make a chemical reaction more favorable and decrease its activation energy (Nelson & Cox, 2008). Enzymes tend to bind with some substrates more than others due to highly favorable conformation and binding interactions; these substrates that bind well can potentially have a high specificity for the enzyme. Specificity allows the enzymes to preferably bind to certain substrates over those competing for the active site (Nelson & Cox, 2008). Figure 1 shows substrates of several shapes and enzymes whose actives sites differ in shape. The blue substrate and enzyme serve as a simplified example of how corresponding shapes can allow for specificity between them. Figure 1, however, oversimplifies the interaction between an enzyme and a substrate, as an enzyme-specific substrate may not fit perfectly into the binding pocket initially (as is indicated by the matching shapes in the figure), but may cause the enzyme to undergo a conformational

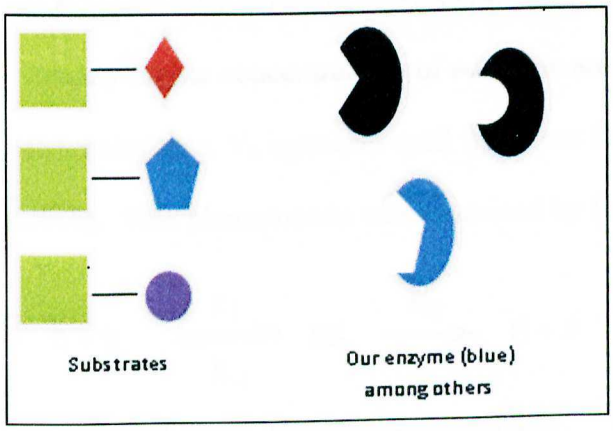

Figure 1. Selection of the correct substrate based on the best fit and chemical characteristics

change that favorable fits the substrate. The enzyme obtains its best substrate interaction when it has an active site that conforms to the reaction transition state; this is the point at which molecules have an equal chance of either returning to substrate or becoming products (Nelson & Cox, 2008). The optimal interaction between an enzyme and its transition state contributes to the specificity of a substrate to a particular enzyme.

By studying the mechanisms of enzyme-substrate reactions, the affinity between a particular substrate and enzyme can be examined. The steady state rate at which the enzyme initially converts substrate(s) into product(s) is called initial velocity (V_0); when steadily higher concentrations of substrate are used against constant enzyme concentrations, V_0 increases until it reaches its maximum value, V_{max} (Nelson & Cox, 2008). This phenomenon was examined by Leonor Michaelis and Maud Menten who developed an enzyme theory and the well-known Michaelis-Menten equation (Menten & Michaelis, 1913). Their enzyme theory was based on the enzyme-substrate reaction shown in Figure 2. Initially, the enzyme and substrate combine in a reversible reaction at a combination rate of k_1 to become enzyme-substrate complex (ES) with a corresponding degradation rate of k_{-1}. The ES complex can then undergo a chemical reaction that results in a separate enzyme and product (P) at a rate of k_2. Michaelis and Menten theorized that upon first combining the enzyme with abundant substrate, the system exhibits steady-state behavior where product is created at a constant rate and [ES] is constant; this initial steady-state behavior is represented by the V_0 value and is the basis for the Michaelis-Menten equation shown below (Nelson & Cox, 2008):

Figure 2. Michaelis -Menten enzyme-substrate reaction (Nelson & Cox 2008)

$$V_0 = \frac{V_{max}[S]}{K_m + [S]} \qquad K_m = \frac{(k_{-1} + k_2)}{k_1}$$

Equation 1. Michaelis-Menten equation and Michaelis constant (K_m) definition

The Michaelis constant (K_m) is commonly used as a measure of affinity between the enzyme and the substrate, although its meaning can be changed based on the rate-limiting steps of the catalysis. Another rate constant, called either the turnover rate or k_{cat}, was defined in order to represent rate limitations caused chemically by reactions at saturation (Nelson & Cox, 2008). The k_{cat} constant is calculated by dividing V_{max} by the total enzyme concentration, and it represents the number of substrate molecules that the enzyme is able to convert to product during a certain period of time. Neither K_M nor k_{cat} provide an appropriate rate measurement that accounts for all aspects of the reaction, so a parameter, called the specificity constant (Equation 2), was developed in order to allow for comparison of turnover between different substrates against the same enzyme.

$$Specificity\ Constant = \frac{k_{cat}}{K_m}$$

Equation 2. Rate of conversion from E+S to E+P called the specificity constant

For the substrate comparisons being performed in this thesis, the specificity constant is used as the main comparative measure between the substrates used in enzymatic testing.

Lipases and Esterases

M. tuberculosis depends on the function of hydrolases in order to survive, grow, and spread throughout a host body. Hydrolases are enzymes that are able to divide a substrate into two separate molecules by hydrolysis. The enzyme breaks a chemical bond in the substrate, replacing the two ends with either an alcohol group (-OH) or hydrogen (-H). By creating smaller and smaller molecules *via* hydrolysis, hydrolases breakdown intra- and extracellular host lipids into useful energy for the pathogen (Cotes *et al.*, 2008). Lipases and esterases are two specific types of hydrolases that are used for this critical lipid metabolism within *M. tuberculosis*.

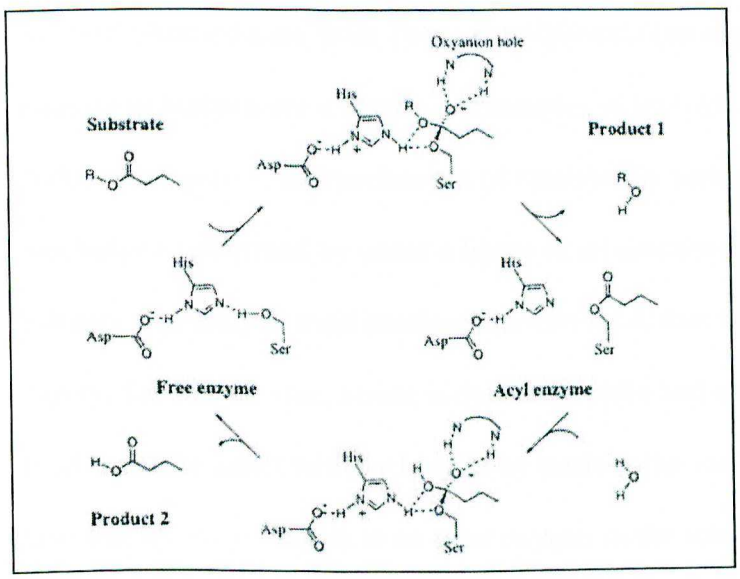

Figure 3. Catal ytic mechanism of serine hydrolase. This mechanism is the same for both lipases and esterases in which an acyl enzyme intermediate is involved in the hydrolytic cleavage that releases one of the cleaved products (Copied from Holmquist, 2000)

A lipase's key function is to break down lipids into fatty acids and alcohols through the hydrolysis of an ester bond (Holmquist, 2000). Biologically, lipases function as a means to hydrolytically break down large lipids within an organism to provide energy and materials for growth (Holmquist, 2000). In humans specifically, lipases are used within adipocytes to break down triacylglycerols and cholesterol; these products are used to maintain the energy balance in the body (Schicher *et al.*, 2010).

Esterases share many characteristics with lipases, including the hydrolysis of ester bonds using the same protein fold (α/β- hydrolase protein fold), the same catalytic triad (nucleophile-His-acid) to perform catalysis, and the assistance in lipid metabolism in biological systems (Holmquist, 2000). The main difference between lipases and esterases is that esterases prefer to cleave at ester bonds located on shorter substrates (<10 carbons) whereas lipases normally cleave at ester bonds on longer substrates (>10 carbons). Acetylcholine esterase from *Torpedo californica* (an electric ray) and butyrylcholine esterase in humans are a couple of examples of biologically relevant esterases (Holmquist, 2000). In Figure 3, the mechanism of catalysis by serine hydrolase, an example of a mechanism performed by either a lipase or an esterase, is shown. Serine hydrolase exhibits the catalytic triad (nucleophile-His-acid) that is present in the α/β- hydrolase family; for this enzyme, serine is the nucleophile and aspartate is the acid. The catalytic triad is able to assist with hydrolysis by positioning the ester correctly; the hydrogen on histidine hydrogen-bonds to an ester oxygen in the substrate while the double-bonded ester oxygen donates electrons that allow for attachment to the serine group via the ester carbon. This process forms a negatively charged tetrahedral transition state that is stabilized by other nearby amino acids *via* the formation of an oxyanion hole (Holmquist, 2000). This transition state allows electrons to be redistributed within the nucleophile-substrate complex such that the R-group on the initial ester substrate is removed; this leaves an acyl enzyme intermediate with part of the original substrate still attached. The addition of water releases a carboxylic acid derivative of the initial substrate, and returns the enzyme back to its free state. Other enzymes in the α/β- hydrolase family perform this catalysis mechanism using their own catalytic triads (Holmquist, 2000).

In TB, specific hydrolytic enzymes, including key metabolic lipases and esterases, have been directly linked to the virulence of *M. tuberculosis* (Singh *et al.*, 2010). Their key roles in the virulence of TB are tied to the ability of TB lipases and esterases to break down host tissues and to utilize the products for energy and growth (Singh *et al.*, 2010). *M. tuberculosis* also stores large quantities of lipids during the dormant phase; this storage lipid can then be broken down by hydrolytic enzymes upon cell reactivation in order to provide the energy to multiply and trigger an infection (Singh *et al.*, 2010). These enzymes also use lipids from host membranes, causing destruction of host cells and increased virulence (Cotes *et al.*, 2008). At least 250 enzymes have been found to be involved in lipid metabolism processes within *M. tuberculosis* (Cole *et al.*, 1998); of these, 24 putative lipases have been identified to assist in both survival during dormancy and host cell infection (Singh *et al.*, 2010). A recent study showed that a lipase within the cell wall of *M. tuberculosis* (RV3802c) has a negative effect on TB cell growth when its catalytic activity is inhibited (West *et al.*, 2011). Without the lipases and esterases used in lipid metabolism, *M. tuberculosis* would be unable to obtain useful energy during its dormant phase, and would likely die within the body before the pathogen could become virulent. Even if the organism was able to survive through dormancy, its ability to infect

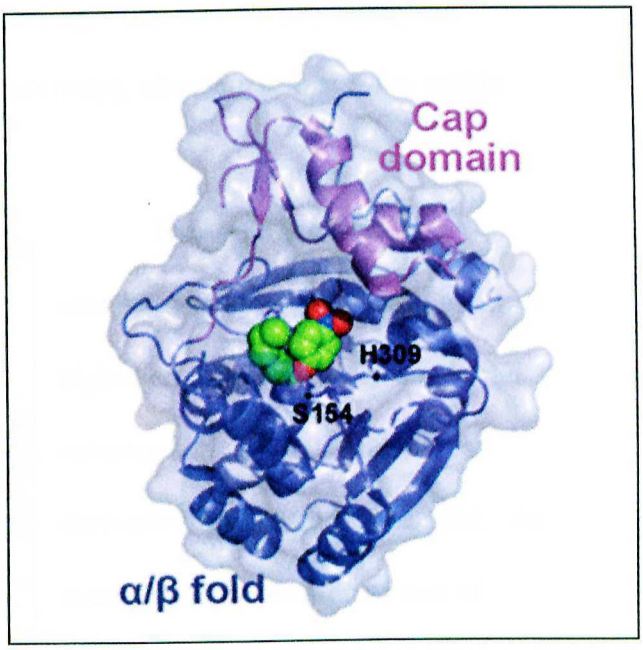

Figure 4. Structure of RV0045c. The α/β fold is shown in dark blue and the cap domain is shown in purple. The binding pocket is highlighted in green and red (Zheng, 2011). Image created in PyMOL (DeLano Scientific, LLC)

host cells would be greatly diminished without the hydrolytic enzymes responsible for membrane breakdown.

For this thesis, substrate testing was performed using the TB hydrolase Rv0045c, shown in Figure 4. Rv0445C is proposed to be responsible for ester and lipid metabolism in *M. tuberculosis*, and recently had its three-dimensional structure determined (Zheng, 2011). Rv0045c is pH and thermally stable under ranges that include physiological conditions (pH 6.0 – 10 and less than 40 °C respectively), allowing it to survive within human host cells (Guo *et al.*, 2010).

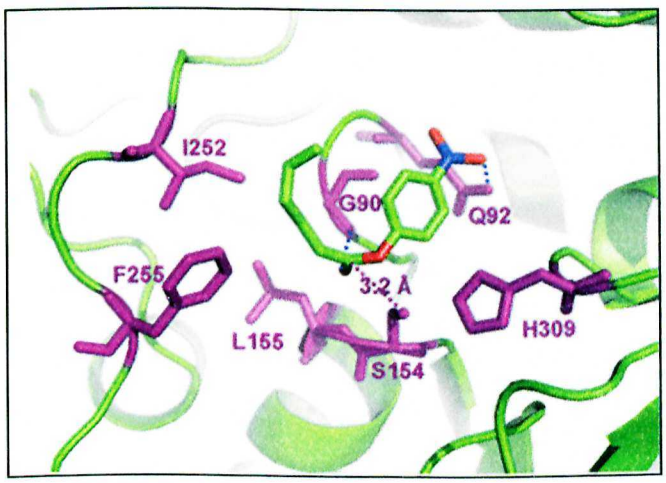

Figure 5. Labeled binding site and active site of RV0045c with *p*-nitrophenyl caproate bound. The binding site is composed of G90, Q92, L155, I252, and F255, and the active site contains S154 and H309. The *p*-nitrophenylcaproate substrate is shown in green, red, and blue colors (copied from Zheng *et al.*, 2011).

Rv0045c is currently classified as an esterase, as it exhibits higher enzymatic activity on shorter esterase substrates (*p*-nitrophenyl derivatives of 2-8 carbon length). An example of the binding of one of these derivatives is shown in Figure 5. The substrate *p*-nitrophenylcaproate (six carbons in the hydrocarbon chain) is bound within the active site, and its hydrocarbon chain is complementary to the binding pocket. This indicates that the Rv0045c enzyme is most likely to act specifically with short-chain esters, and based on sequence homology to known bacterial esterases is likely an enzyme involved in the lipid metabolism of *M. tuberculosis* (Zheng *et al.*, 2011). Structures in the *p*-nitrophenylcaproate substrate were considered when

synthesized substrate design was being determined; the designs both incorporate a short chain of comparable length to *p*-nitrophenylcaproate and one design contains a phenyl group similar to those on all of the tested *p*-nitrophenyl derivatives. The details of the chosen substrate designs are discussed later in the background section.

Fluorescence

Fluorescence is the light that is emitted from an excited particle after it has absorbed energy ("Fluorescence", 2012). The Jablonski Diagram in Figure 6 shows that energy jumps from a base level, S_0, to a higher energy level, S_2, upon excitation. The energy of the excited S_2 level is then decreased by internal conversion or nonradiative decay and then the remaining energy is released as fluorescence until energy returns to S_0 (Lavis & Raines, 2008). This fluorescent radiation is often within the visible light spectrum and can be measured using spectroscopy. In addition, fluorescence ceases almost immediately when the energy source responsible for its excitation is removed ("Fluorescence", 2012). These characteristics make fluorescence a valuable tool in biological and chemical measurements in which observation and controlled variables are crucial.

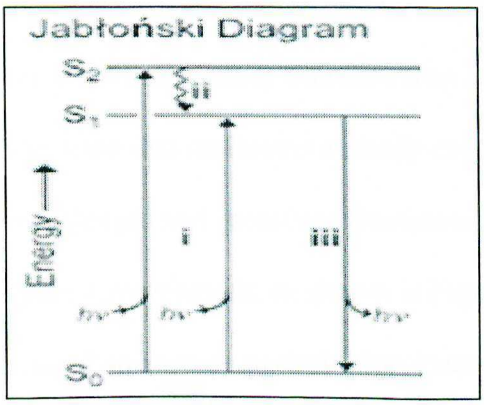

Figure 6. Ja blonski Diagram showing the energy levels of a photon in its excited state (i), after its internal conversion (ii), and during its fluorescence (iii). (copied from Lavis & Raines, 2008).

Fluorescence is found in natural biological settings. For instance, fluorescence has been observed in a type of wood called *Lignum nephriticum*, in the chlorophyll found in plants, in various types of fluorite crystals, and in the jellyfish *Aequorea victoria* (Valeur & Berberan-Santos, 2011). Researchers have taken these observations and developed the fluorescent molecules or fluorophores found in nature and new synthetic fluorophores for laboratory use (Nelson & Cox, 2008).

Fluorescence is typically measured by a fluorimeter that first stimulates a sample with light (excitation) before recording the emission spectra of the molecule in arbitrary units (Jameson *et al.*, 2003). This measurement can create two types of fluorescence spectra. The first is an emission spectrum that maintains a steady excitation wavelength and measures fluorescence intensity against wavelength, as shown in Figure 7. The second spectrum, an excitation spectrum, measures the fluorescence intensity at a specified emission wavelength; this measure reveals which wavelength of light is most efficient at exciting the molecule. A molecule's fluorescence can be used in a number of biological and chemical experiments, including distance estimations and the determination of subunits in a particular structure (Jameson *et al.*, 2003). Fluorophores such as anthracene for the indication of ions, quinolium for the presence of chloride ion, and indole for sensing calcium, have become important tools for researchers studying a variety of biological signals (Lavis & Raines, 2008). In this thesis, the

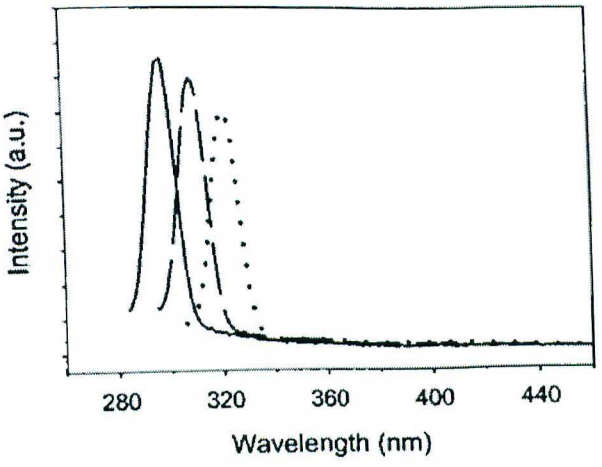

Figure 8. Examples of emission scans. Fluorescence intensity peaks represent emission scan of 50mM phosphate buffer at 270 nm (solid), 280 nm(dashed), and 290 nm (dotted) (copied from Jameson *et al.*, 2003)

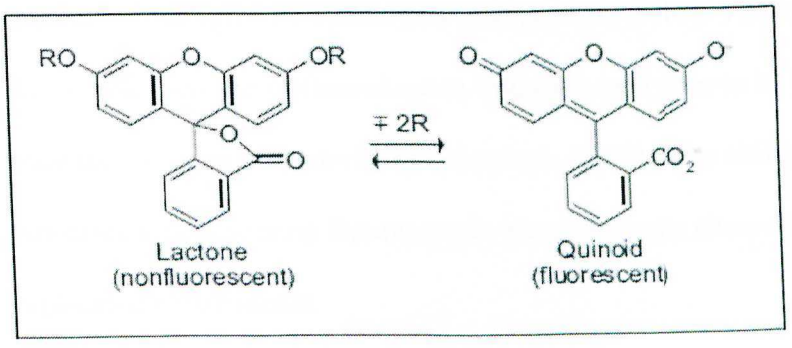

Figure 7. Transition of fluorescein between its lactone and quinoid state based on attachment of R groups (copied from Lavis & Raines, 2008)

fluorophore, fluorescein (Jameson *et al.*, 2003), is used to create latent fluorogenic ester substrates for fluorescence studies on enzyme-substrate interactions. Fluorescein is a near perfect emitter in its standard quinoid form, but when it is substituted at its phenolic oxygen, it undergoes a tautomerization to create the lactone form, which is nonfluorescent (shown in Figure 8). In this thesis, an ester will be attached as the R-groups on fluorescein to lock it in the nonfluorescent lactone form, then addition of an esterase, such as RV0045C, will hydrolyze the ester, releasing the highly fluorescent fluorescein (Lavis & Raines, 2008). Because the substrates are protected in the lactone state and fluoresce only when acted upon by an enzyme, we call them latent fluorophores. One way to create latent fluorophores is through the addition of acetyl groups to the phenolic oxygens on fluorescein *via* an acylating agent; the resulting substrate is fluorescein diacetate (Lavis *et al.*, 2010). Although the creation of ester groups prevents the fluorescence of this substrate, the overall stability of fluorescein diacetate is low; this is caused by low pK_a values that make the ester bonds more likely to break in aqueous solution (Lavis *et al.*, 2010). To prevent substrate breakdown in aqueous solution, acetoxymethyl (AM) ethers are used to mask the ester group and significantly decrease the rate of background hydrolysis (Lavis *et al.*, 2010). The low background hydrolysis allows the increase in fluorescence intensity over time to be only an indication of the esterase's enzymatic activity (Lavis *et al.*, 2010). The ability to quantify enzyme-substrate kinetics using fluorescence measurements allows for reliable and easily replicated experiments.

Substrate Synthesis

For my research, I chose two new latent fluorophores to synthesize. These two substrates were designed to examine the importance of substrate flexibility to hydrolysis by Rv0045c. Limited work on the substrate specificity of Rv0045c suggested that it preferentially hydrolyzed substrates with shorter carbon chains (< 10 carbons) (Guo *et al.*, 2010). The carboxylic acids used to synthesize the fluorogenic substrates are 2-phenylbutyric acid (a) and 2-ethylbutyric acid (b) (Figure 9). These acids contain nine carbons and six carbons, respectively, and thus fit the previously established substrate criteria. These chains on both carboxylic acids are completely composed of hydrocarbons because the Leucine155, Ile252, and Phe255 in the active site, shown in Figure 5, are hydrophobic (Zheng *et al.*, 2011). The carbon ring in 2-phenylbutyric acid, Figure 9a, increases the steric property of the substrate whereas the two ethyl groups in 2-ethylbutyric acid, Figure 9b, allow for flexibility at the tertiary carbon. The comparison of substrate hydrolysis of these two substrates by Rv0045c will allow a comparison of similar substrates with different steric properties.

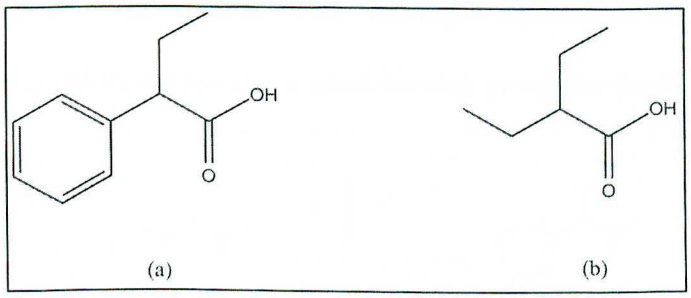

Figure 9. The initial carboxylic acids: 2-phenylbutyric acid (a) and 2-ethybutyric acid(b).

The synthetic method used to create the desired substrates was modeled after the synthesis of Tian and colleagues (Tian *et al.*, 2010). The synthetic method contains three main steps (Figure 10): (1) chloromethylation of the carboxylic acid, (2) iodination *via* an Sn2 reaction to replace the chlorine atom, and (3) the combination of the iodomethyl

ester with fluorescein to produce the final product (Tian *et al.*, 2010). The alkylation at the carboxylic acid (Figure 10 i) was achieved using chloromethyl chlorosulfate, a common chloromethylating agent (Power *et al.*, 2004). The addition of the chloromethyl group greatly increases the hydrolytic stability of the of this intermediate substrate and the chlorine provides a good leaving group for the following step. The second step, Figure 10 ii(1), in the synthesis then replaces the chlorine atom with an iodine atom and was performed by a classic Sn2 reaction, known as the Finkelstein reaction (Finkelstein, 1910). In the Finkelstein reaction, acetone is used as the reaction solution because the reactant sodium iodide is soluble in acetone, but the product sodium chloride is insoluble in acetone. Thus, as the reaction proceeds, the iodomethylated ester product will remain in solution while the sodium chloride will precipitate out of solution and drive the reaction (Finkelstein, 1910). In the third and final synthetic step (Figure 10 ii(2)), the iodomethylated ester and fluorescein were combined in the presence of silver oxide. Depending on the electrophile in the reaction, this reaction can produce monoalkylated, monoacylated, or hybrid monoalkylated and monoacylated fluorescein,

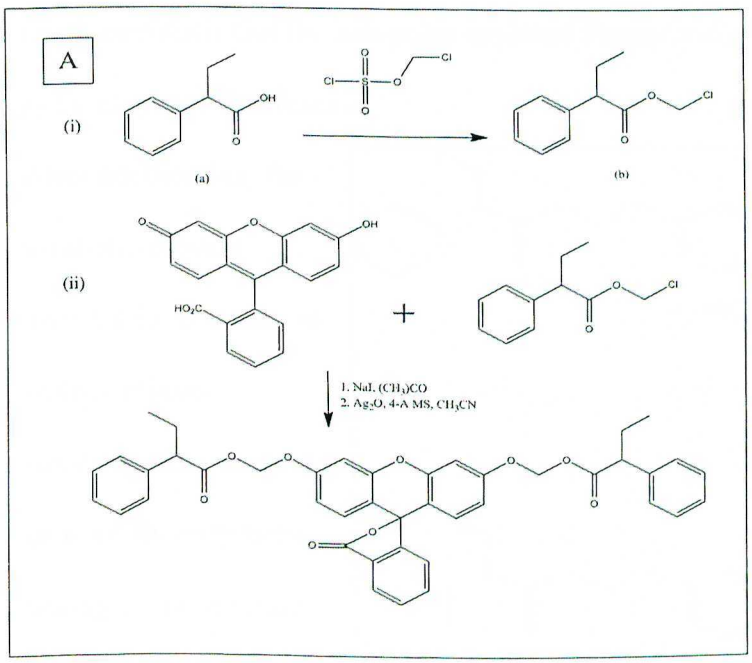

Figure 10. Syn thetic schematic of substrate A, fluorescein di(phenylbutyloxymethyl ether). Step one (i) chloromethylated the carboxylic acid, and steps two and three (ii1 and ii2) allow fluorescein to combine with the acyloxymethyl group (Tian *et al.*, 2010).

where the dialkylated products are desired over the diacylated or hybrid products due to their higher stability. The reaction conditions for this final step favored alkylation because the combination of methyl iodide (iodide is in the iodomethylated ester) and silver oxide is commonly used for N-methylation which attaches a methyl carbon to a single-bonded oxygen (Hughes, 2009). A column purification of the final solution was conducted such that the less polar acylated fluorescein could be separated from the more polar alkylated fluorescein.

After purification, the synthetic process allowed for production of two unique substrates that could be used for enzymatic testing of the esterase Rv0045c (Figure 11). These substrate products will hitherto be referred to as Substrate A and Substrate B for simplification and clarity.

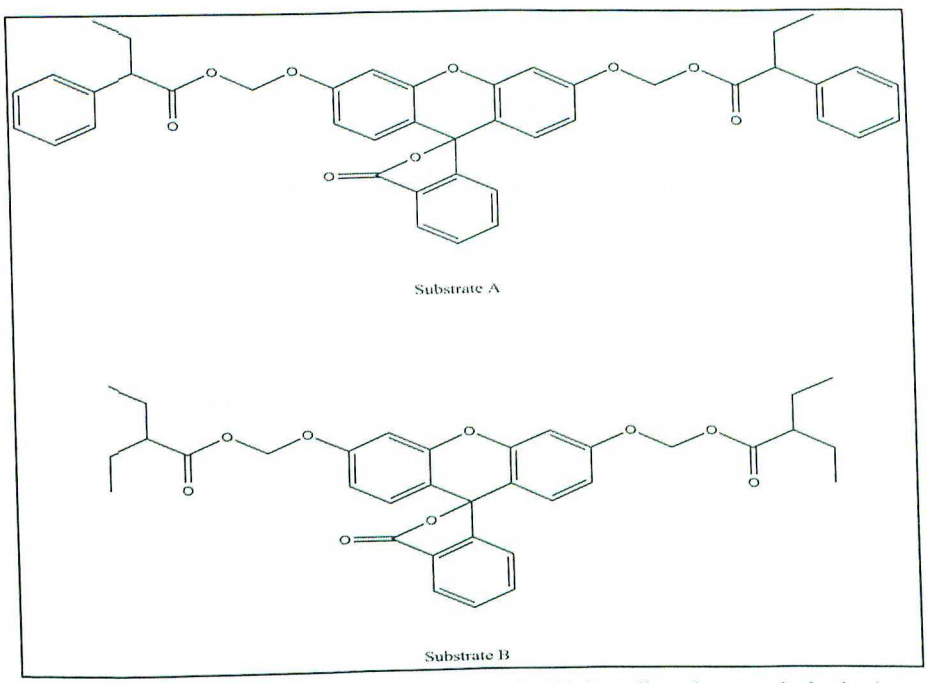

Figure 11. Final s ynthetic products: fluorescein di(phenylbutyloxymethyl ether), Substrate A, and fluorescein di(ethylbutyloxmethyl ether), Substrate B.

Treatment for Dormant Tuberculosis (TB)

The goal of this research was to synthesize two novel latent fluorophores, to understand the substrate specificity of an essential TB esterase, and to use this information to design potential inhibitors of Rv0045c. Because TB lipases and esterases aid in survival during the dormant phase, eliminating these important enzymes could either reduce the risk of developing active TB or eradicate *M. tuberculosis* completely (Gurdyal *et al.*, 2010). A potential treatment option for dormant TB is to use prodrugs. Potential prodrugs would be substrate-antibiotic complexes connected together by an ester bond to specifically target *M. tuberculosis*. This ester bond would act as a protecting group to the antibacterial, allowing it to only release upon hydrolysis by an esterase that recognizes this protecting group. To design an effective prodrug, requires significant information about substrates or protecting groups with high activity and selectivity for Rv0045c. This thesis provides an initial step in understanding the specificity and selectivity of Rv0045c. Results from this thesis can then be refined through the synthesis of new substrates with potentially higher activity and specificity until a novel treatment for dormant TB infections can be developed.

Results and Discussion

The three synthetic steps described in the background and materials and methods sections of this thesis were used to create two separate fluorophores, fluorescein di(phenylbutyloxymethyl ether), Substrate A, and fluorescein di(ethylbutyloxmethyl ether), Substrate B. The first step in the synthesis was to chloromethylate the carboxylic acid starting material: 2-phenylbutyric acid (a) and 2-ethylbutyric acid (b) (Figure 9). The product of this reaction then underwent the Finkelstein reaction in the second synthetic step in order to substitute the chlorine ion with an iodine ion, creating an iodomethylated ester product. The products from the first and second synthetic steps were not characterized by liquid chromatography mass spectroscopy (LCMS) or proton NMR, as the product of these reactions were immediately used in the next step in the synthesis. The final synthetic step combined fluorescein with the iodomethylated ester in the presence of silver oxide. In this work, acylation occurred when the phenolic oxygen acted as a nucleophile to the electrophilic ester carbon in the iodomethylated ester product. The attachment at this location caused the detachment of the iodomethylated oxygen such that the final attachment contained no methylene carbon ester. Additionally, alkylation occurred when phenolic oxygen on fluorescein acted as a nucleophile to the electrophilic methylene carbon; the oxygen's electrons formed a bond with the carbon and released the iodine atom. Since fluorescein contains two phenolic oxygens, it can produce a diacylated, dialkylated, or hybrid monoacylated and monoalkylated product. These three products are shown in Figure 12a and Figure 13a for Substrate A and Substrate B, respectively. The synthesis also had the ability to produce monoacylated

and monoalkylated products in which only one of the phenolic oxygens on fluorescein was acylated or alkylated (Figure 12b, Figure 13b).

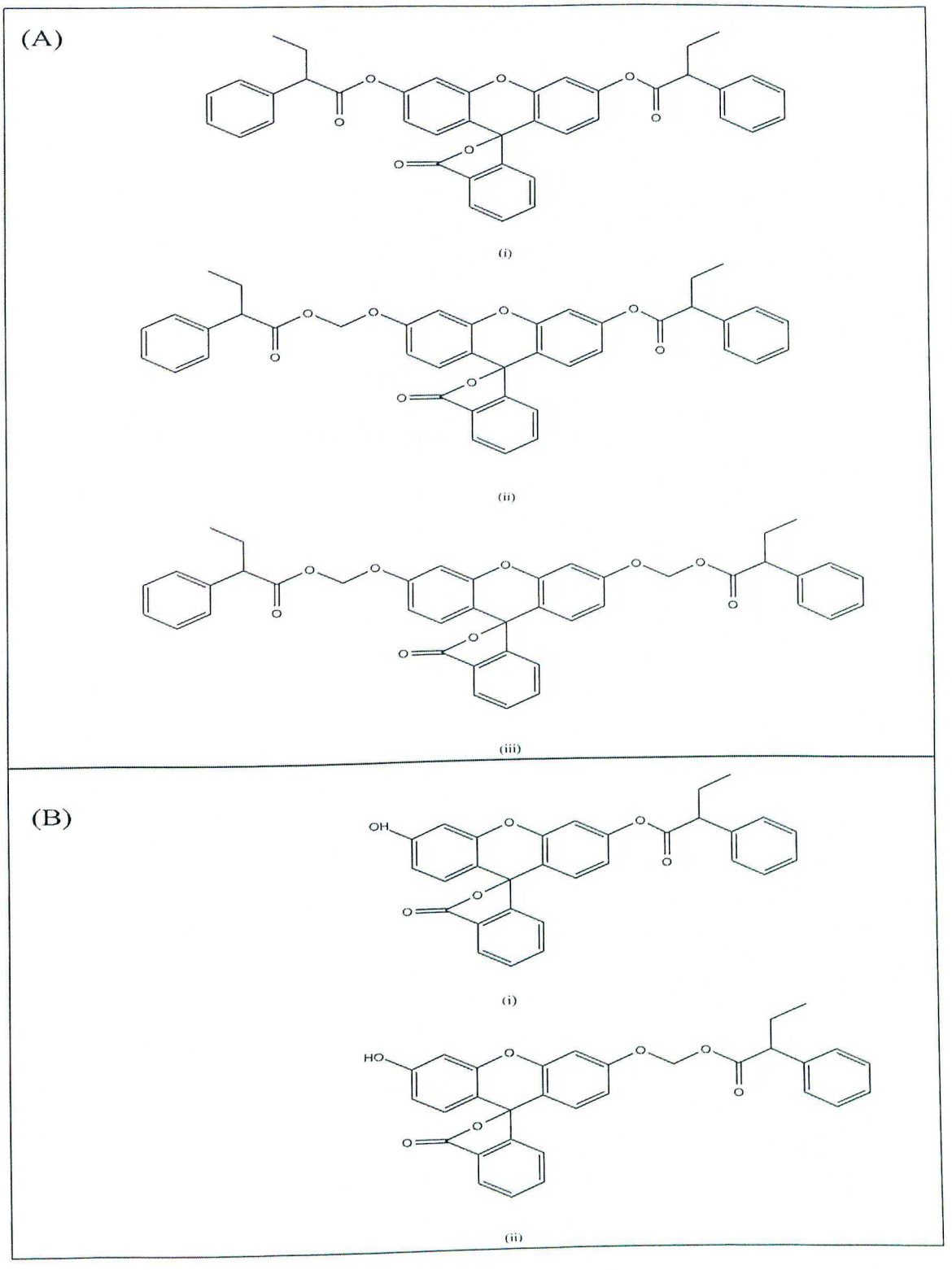

Figure 12. Potential products of fluorescein di(phenylbutyloxymethyl ether), or Substrate A. Products with two attachments (A) are diacylated (i), hybrid monoacylated and monoalkylated (ii), or dialkylated (iii). Products with only one attachment (B) are either monoacylated (i) or monoalkylated (ii).

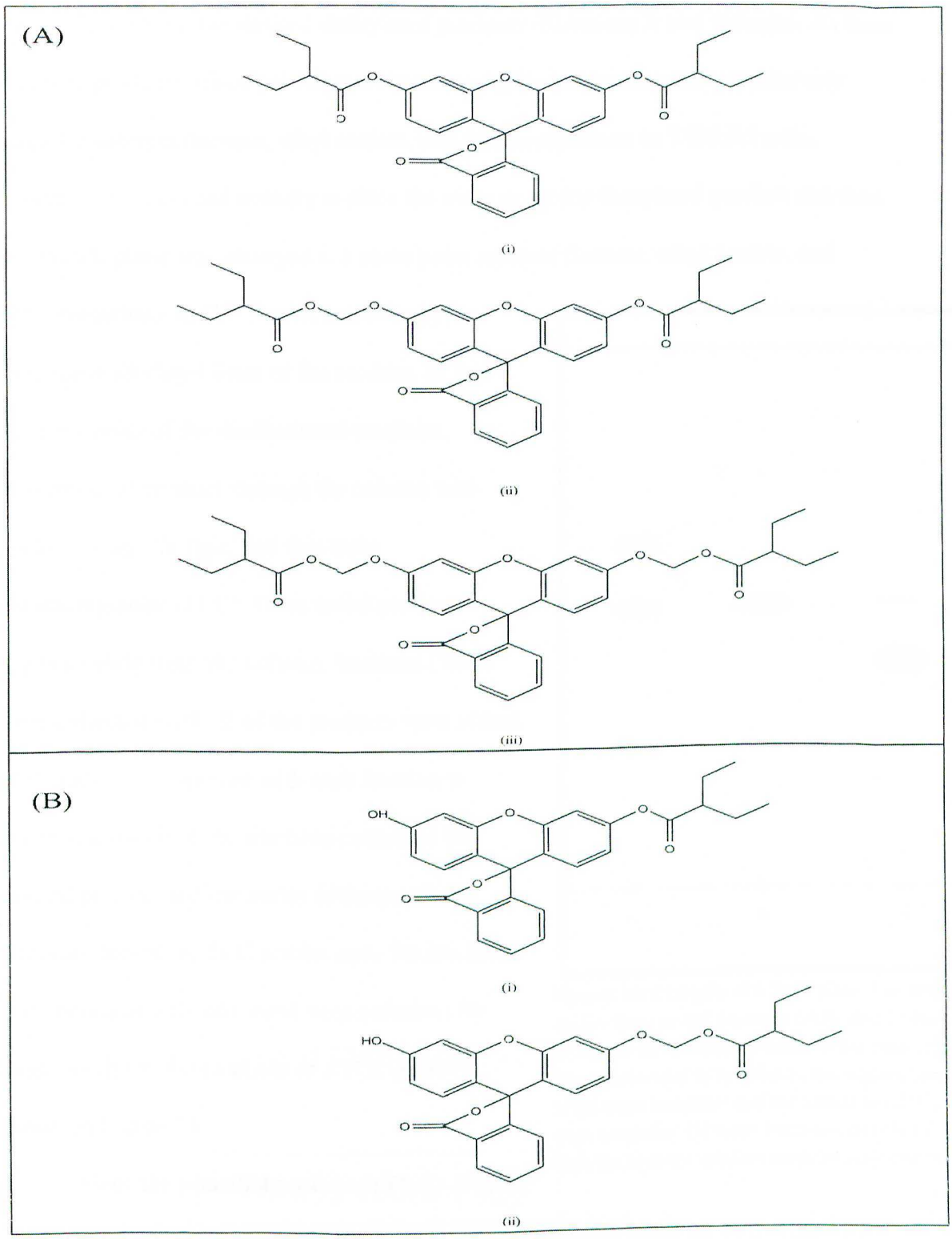

Figure 13. Potential products of fluorescein di(ethylbutyloxymethyl ether), or Substrate B. Products with two attachments (A) are diacylated (i), hybrid monoacylated and monoalkylated (ii), or dialkylated (iii). Products with only one attachment (B) are either monoacylated (i) or monoalkylated (ii).

To separate the desired dialkylated products (Substrate A and Substrate B) from the other products, silica gel column chromatography was performed. A relatively nonpolar solution (hexane, ethyl acetate, and dichloromethane in 7.5:0.5:2 ratio, respectively) was used initially to elute the more nonpolar diacylated product and then the mobile phase was changed to a more polar mixture (hexane, ethyl acetate, and dichloromethane in 7:1:2 ratio, respectively) to elute the dialkylated form of the product, as it is the most polar of the disubstituted products. Movement of product through the column was tracked using UV light and thin layer chromatography (TLC). Once initial products began to elute from the column, fractions (5mL) were collected until all of the products were eluted. TLC plates were spotted with each fraction to determine which of the fractions contained the desired product and the purity of the purified fractions. Based on TLC results, only the fractions that contained only one band were collected for future analysis. An example of a TLC plate is shown in Figure 14.

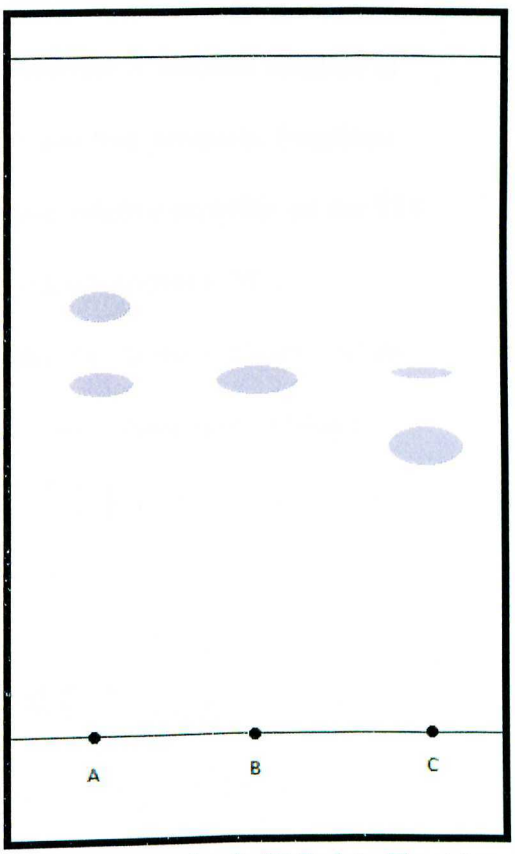

Figure 14. Example of a TLC plate. The strips shows three tested fractions (A,B, and C) that have produced three distinct bands. In the case of the TLC performed in this thesis, the highest band (A) is the most nonpolar and the lowest band (C) is the least nonpolar. Of these fractions, only B would be kept because its solution contains only one product.

Since the potential products all have slightly different polarities, band separation was observable between them by TLC. Since the solution used for TLC had a high ratio of hexane (nonpolar), the most nonpolar

compounds traveled farther on the plate than the polar compounds. The desired dialkylated product is the most polar of the disubstituted products, so it was expected to elute last from the silica column and to be the lowest band observed on the TLC. Products that were either monosubstituted or nonsubstituted are even more polar than the dialkylated products; for this reason, it was assumed that these were left on the column after band separation and that they would not be observed on the TLC plates.

Analysis of the TLC results indicated that the Substrate A reaction resulted in three distinct products, and that the Substrate B reaction had four products. Fractions from the silica gel column containing products of identical relative mobility on the TLC were combined to create the samples used for product characterization. This characterization utilized two laboratory tools to determine the chemical identity of the collected products (described more later). The synthesis and column purification of Substrate A provided three separate products, and the synthesis of Substrate B produced four products.

The seven total products from substrate A and B were then characterized by liquid chromatography mass spectrometry (LCMS) and proton nuclear magnetic resonance spectroscopy (H NMR). The LCMS was conducted under positive ionization mode and identified the molecular weights (MWs) of any parent ions in solution and their relative concentrations. An observed MW close to the expected MW of the desired dialkylated product was considered strong evidence that the correct product was synthesized. In addition, H NMR spectroscopy was performed to determine if the proton signals matched those that would be expected based on product predictions from the LCMS. Due to H NMR contamination, the following peaks should be disregarded due to the contaminant

solution they are known to represent: 7.26ppm ($CDCL_3$, s), 5.30ppm (dichloromethane, s), 4.12ppm (ethyl acetate, q), and 2.17ppm (acetone) (Gottlieb, H. 1997). The figure legends of each H NMR spectrum give peak integration details for those that are significant in determining product structure.

Substrate A

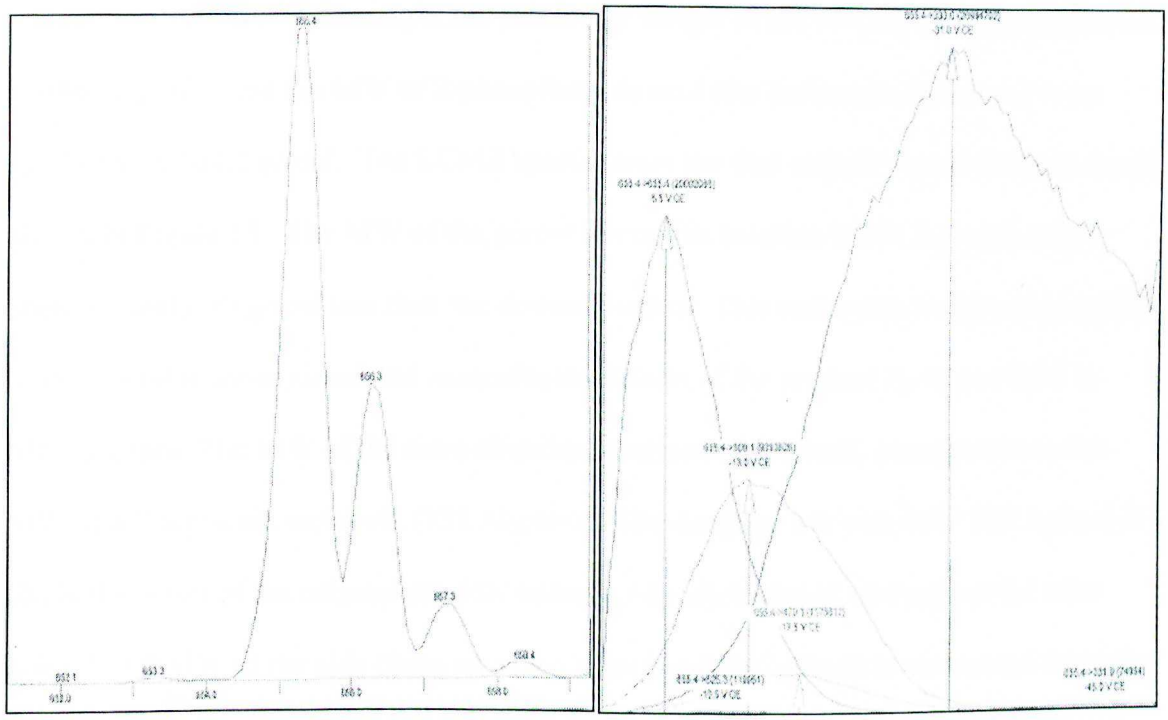

Figure 15. LCMS spectra of Substrate A, sample 1. The MW for parent ions (left) and the MW frequency of daughter fragments (right). The most frequent parent ion MW is 655.4 g/mol and the top three fragment MWs are 333 g/mol, 509.1 g/mol, and 479.1 g/mol

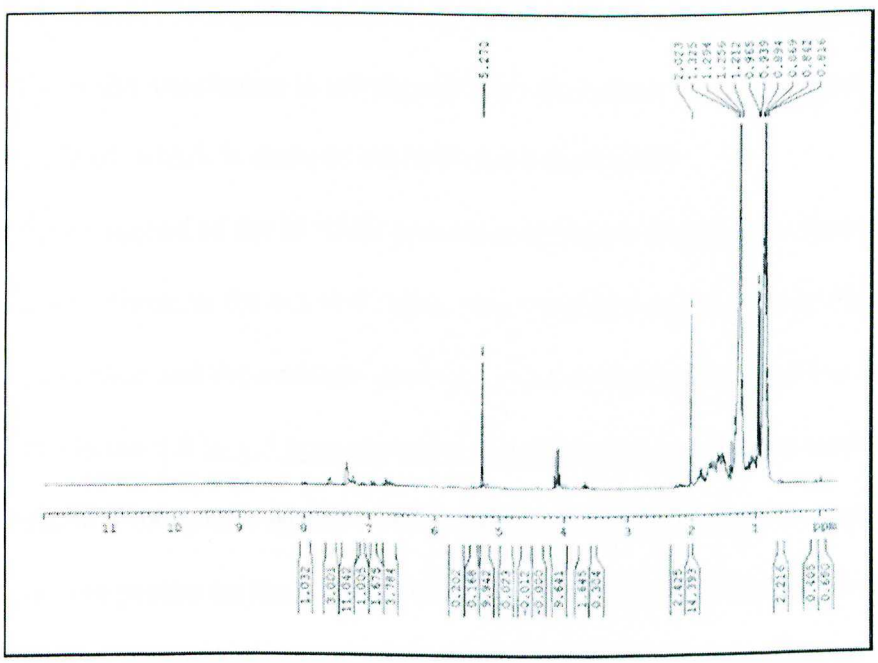

Figure 16. H NMR spectrum for Substrate A, sample 1. 1H NMR (300MHz, $CDCl_3$) δ (ppm): 8.00 (d, 1H), 7.60 (quin, 3H), 7.10 (d, 2H), 6.95 (d, 4H), 6.71 (ddddd, 4H), 3.68 (t, 2H), 2.20 (sep, 2H), 0.95 (d, 4H)

For Substrate A, the expected molecular weight of the desired dialkylated product is 686.71 g/mol, and the MW of 2-phenylbutyric acid (the carboxylic acid used in its synthesis) is 164.2 g/mol. The LCMS spectra from the first sample from Substrate A are shown in Figure 15. The MW of the parent ion in this solution is 655.4 g/mol; this is approximately 30 g/mol less than the desired product. This molecular weight corresponds to the hybrid monoacylated and monoalkylated form of the product A, whose MW is 656.71 g/mol. The MW of the most abundant fragment, 333g/mol, corresponds to the MW of a fluorescein molecule (332.31g/mol). The daughter ion with MW 509.1g/mol is likely the result of the original product losing a side chain due to cleavage at the ester bond. If the MW of the side chain up to the ester bond (147 g/mol) is subtracted from the MW of the parent ion (655.4 g/mol), the weight of the expected molecule would be 508.4 g/mol. The MW 479.1 g/mol fragment is the result of the product losing the entire alkylated attachment (176.2 g/mol), including the ester bond and the methylene carbon. When the MW of the attachment is subtracted from the parent ion MW, the calculated MW is 479.2 g/mol, which is close to the MW from the LCMS.

The figure legend of the H NMR spectrum in Figure 16 gives the locations of prominent peaks. Those in the 6.5 to 8.5ppm range indicate the protons on the phenyl groups in the R-group and the multiple aromatic rings in fluorescein, and the noisy cluster of peaks in the 0.8 to 1.5 ppm represent the alkane protons on the ethyl groups. The triplet peak at 3.68 ppm is specific to the products of the substrate A synthesis; this peak represents the proton on the tertiary carbon between the ethyl and phenyl groups on the attachment. For the H NMR spectrum to indicate the presence of hybrid monoacylated and monoalkylated product, a singlet, 2H peak should be seen at

approximately 5.8ppm. This peak would indicate the presence of one methylene carbon. This peak does not exist, though it may have shifted toward 5.4ppm where it could be masked by the dichloromethane contaminant peak.

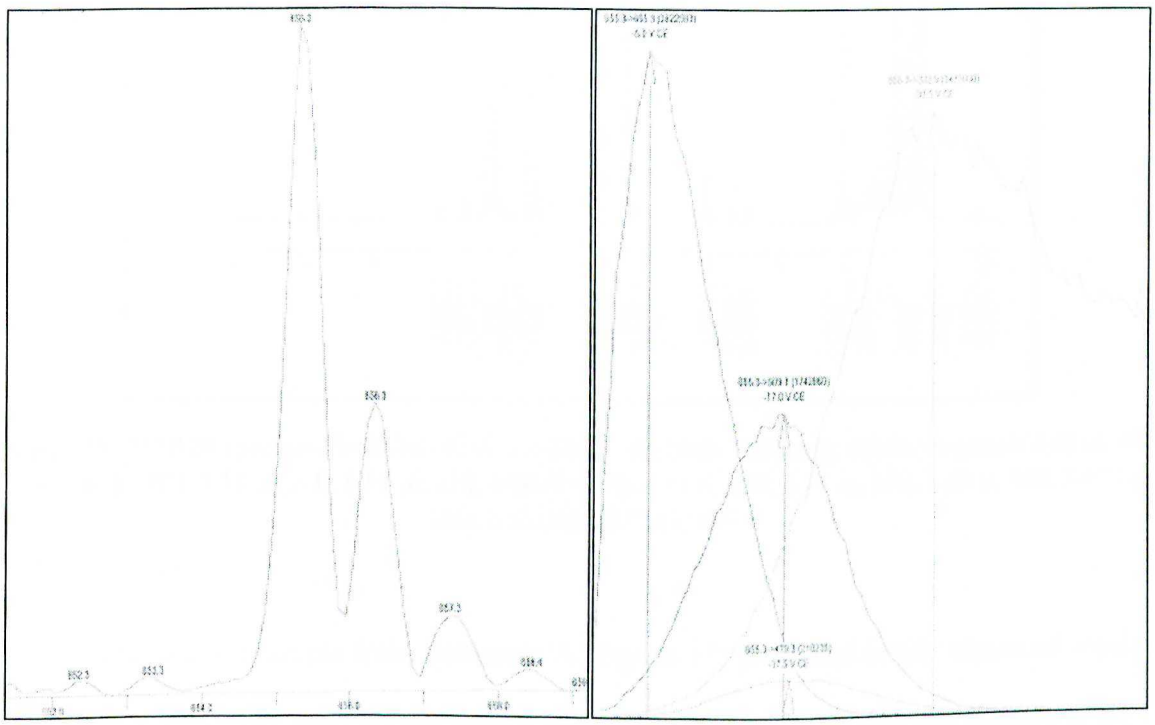

Figure 17. LCMS spectra of Substrate A, sample 2. The MW for parent ions (left) and the MW frequency of daughter fragments (right). The most frequent parent ion MW is 655.3 g/mol and the top three fragment MWs are 332.9 g/mol, 509.1 g/mol, and 479.3 g/mol

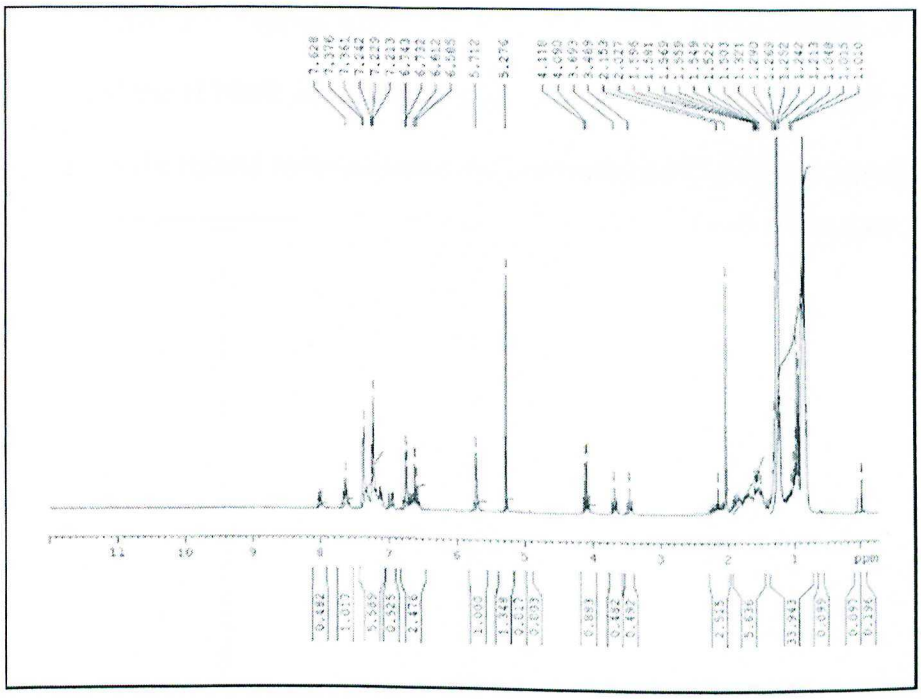

Figure 18. H NMR spectrum for Substrate A, sample 2. 1H NMR (300MHz, $CDCl_3$) δ (ppm): 8.00 (d, 1H), 7.63 (quin, 2H), 7.11 (d, 2H), 6.97 (d, 1H), 6.68 (m, 5H), 5.71 (s, 2H), 4.11 (q, 2H), 3.69 (t, 1H), 3.47 (t, 1H), 2.15 (sep), 0.95 (d, 4H)

The second sample from Substrate A (Figure 17) produced nearly identical results to those from the first sample (Figure 15). Again, the data support that this sample is the hybrid monoacylated and monoalkylated substrate product due to the parent ion MW being 655.3 g/mol, around 30 g/mol lower than the expected MW. The MWs of the daughter ions (332.9g/mol, 509.1g/mol, 479.3g/mol) are nearly identical to those from the first sample (332.31g/mol, 509.1g/mol, 479.1g/mol). They are again presumed to represent the fluorescein molecule, the parent ion without the acylated attachment, and the parent ion without the alkylated attachment, respectively.

The H NMR spectrum in Figure 18 shows aromatic and alkane proton peaks similar to those from the first sample. The two triplets at 3.69ppm and 3.47 ppm represent the tertiary carbon protons specific to this substrate. Unlike the first sample, a

2H singlet peak exists at 5.71ppm, indicating the presence of one methylene carbon. The LCMS spectra and the H NMR spectrum complement each other and strongly indicate that this product is the hybrid monoacylated and monoalkylated substrate product.

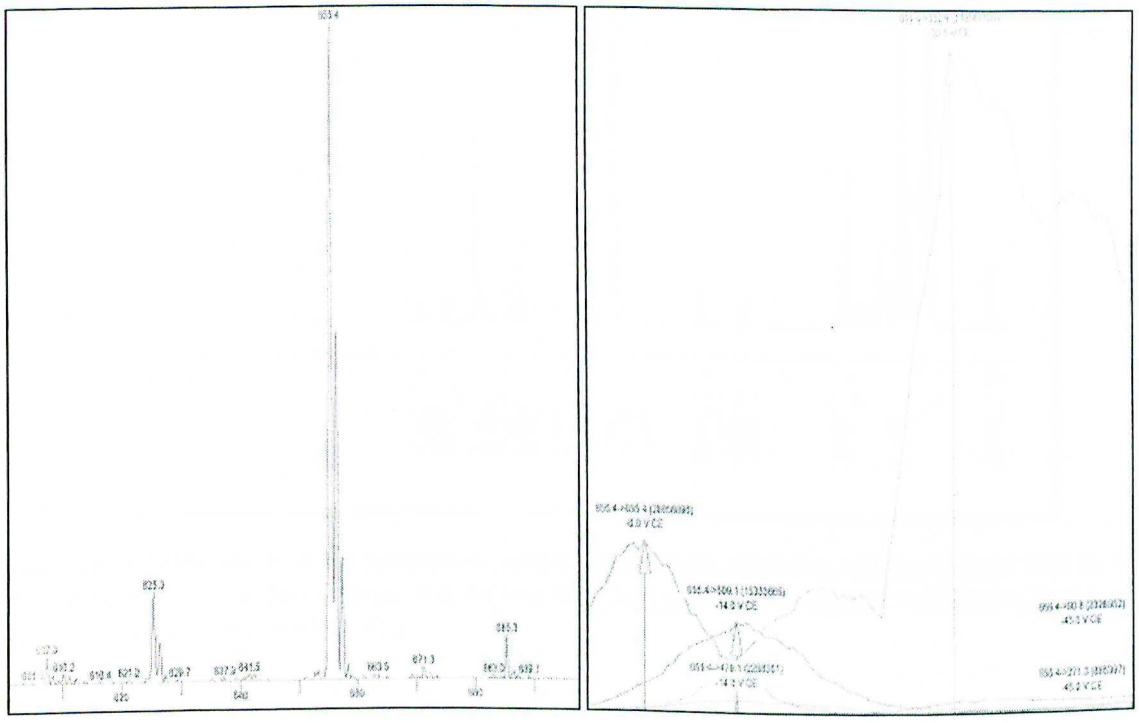

Figure 19. LCMS spectra of Substrate A, sample 3. The MW for parent ions (left) and the MW frequency of daughter fragments (right). The most frequent parent ion MW is 655.4 g/mol and the top three fragment MWs are 332.9 g/mol, 509.1 g/mol, and 479.1 g/mol

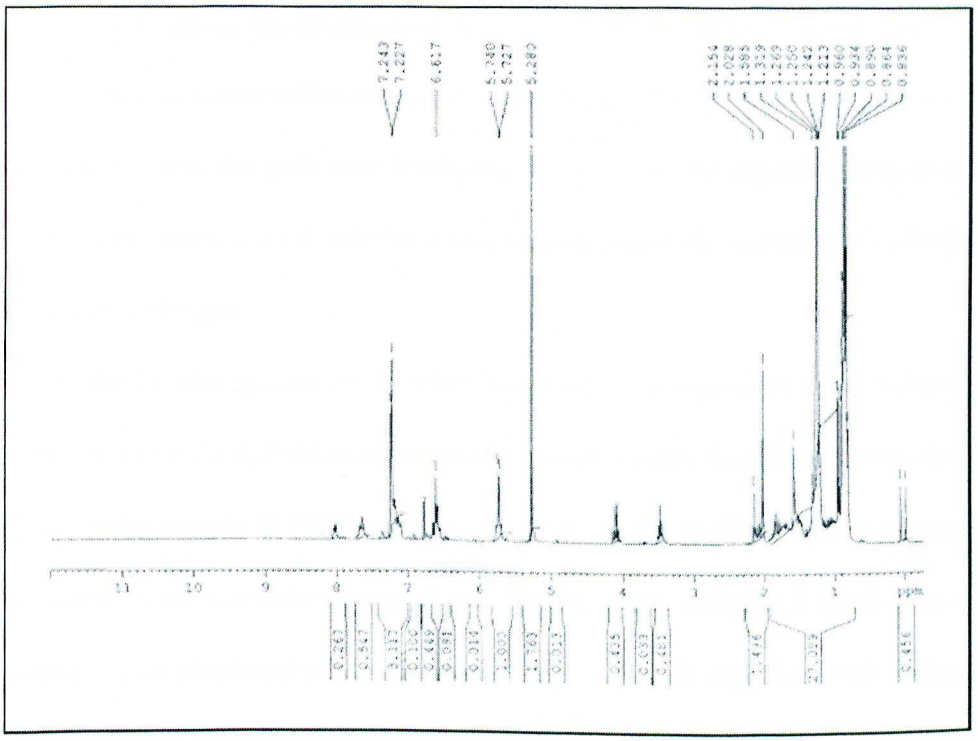

Figure 20. H NMR spectrum for Substrate A, sample 3. 1H NMR (300MHz, $CDCl_3$) δ (ppm): 8.03 (d, 1H), 7.65 (quin, 2H), 7.13 (d, 2H), 6.78 (d, 2H), 6.61 (q, 4H), 5.73 (d, 4H), 4.11 (q, 2H), 3.48 (t, 2H), 2.10 (sep, 6H), 1.95-1.45 (m, 10H), 0.95 (d, 6H)

Identical to samples 1 and 2 from substrate A, sample 3 from substrate A gave a parent ion MW, which indicated that the synthesis produced the hybrid monoacylated and monoalkylated substrate product. Figure 19 shows that the parent ion MW was once again 655.4 g/mol, and 30 g/mol less than the expected MW of Substrate A. The three most frequent daughter fragment MWs (332.9g/mol, 509.1g/mol, 479.1g/mol) are nearly identical to those for the first and second samples; these MWs represent fluorescein, the parent ion without its acylated attachment, and the parent ion without its alkylated attachments, respectively.

Figure 20 shows a doublet peak with 4H in this vicinity (5.73ppm); although it appears in the correct location, this peak does not represent the methylene carbon ether because (1)

the double peak indicates the presence of adjacent hydrogen and (2) the 4H indicates that this occurs in two locations rather than one. In addition, the 5.73ppm peak in the third sample does not match the peak that would be indicative of the desired dialkylated product; if it were, there would only be a singlet peak since the methylene carbon ether has no adjacent hydrogen.

From the LCMS spectra of the three Substrate A samples, all three Substrate A samples appear to be the hybrid monoacylated and monoalkylated form of the desired product. All of the sample H NMR spectra (Figures 16, 18, and 20) also show nearly identical results in the aromatic region (6.5-8.5ppm). Only samples 2 and 3 show the unique feature of an alkylated product, which is a 2H peak at approximately 5.8ppm, indicating the presence of the methylene carbon in the ether (Figure 17). Proper identification by NMR is also complicated by the presence of a doublet peak with 4H in this vicinity (5.73ppm) for sample (Figure 21). This splitting pattern would not be expected for the methylene carbon, which would be a singlet, even in the dialkylated form. Based on this analysis, two items can be confirmed: (1) only the second sample can confidently be labeled as hybrid monoacylated and monoalkylated substrate product and (2) none of the samples are the desired dialkylated Substrate A. The nearly identical and undesired synthesis results from these three samples suggest one of two things: either the synthesis did not produce any of the desired substrate product or the purification in the column was halted before the desired substrate had come off of the column. As explained previously, the desired substrate is the most polar of the possible products and would have been collected from the column last; ending the purification too early would have missed the collection of the desired dialkylated product.

Substrate B

The expected MW of Substrate B is 590.63 g/mol, and 2-ethylbutyric acid, the carboxylic acid used to create the substrate, has a MW of 116.16 g/mol.

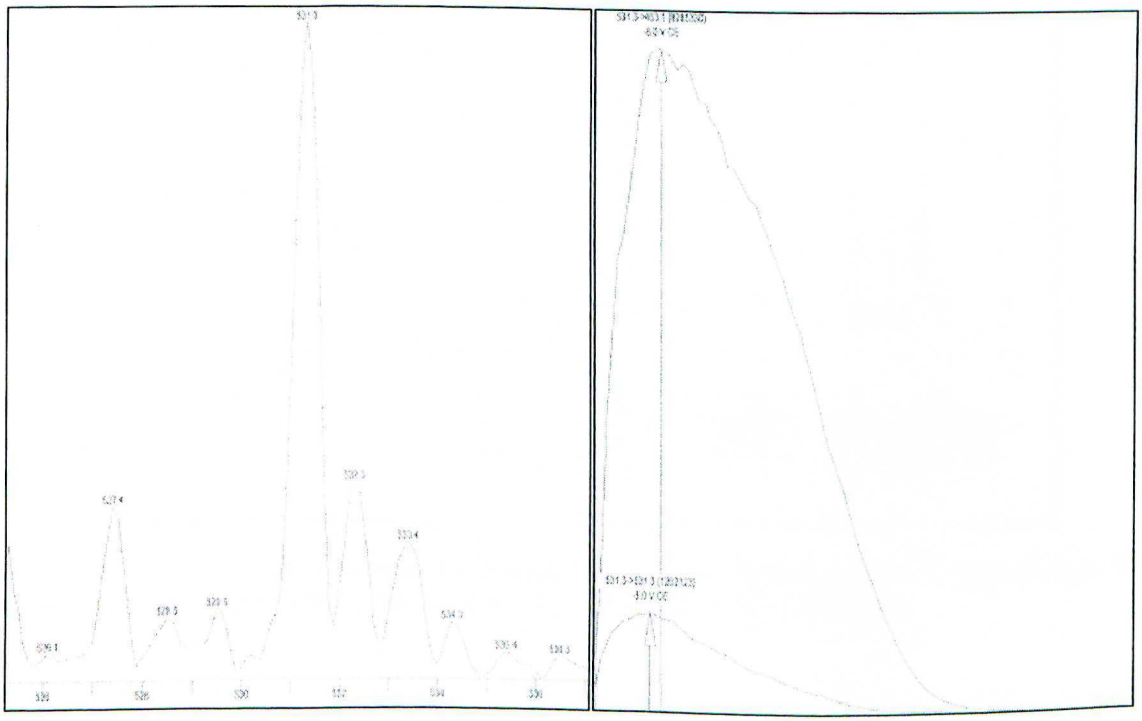

Figure 21. LCMS spectra of Substrate B, sample 1. The MW for parent ions (left) and the MW frequency of daughter fragments (right). The most frequent parent ion MW is 531.3 g/mol and the only fragment MWs is 453.1 g/mol.

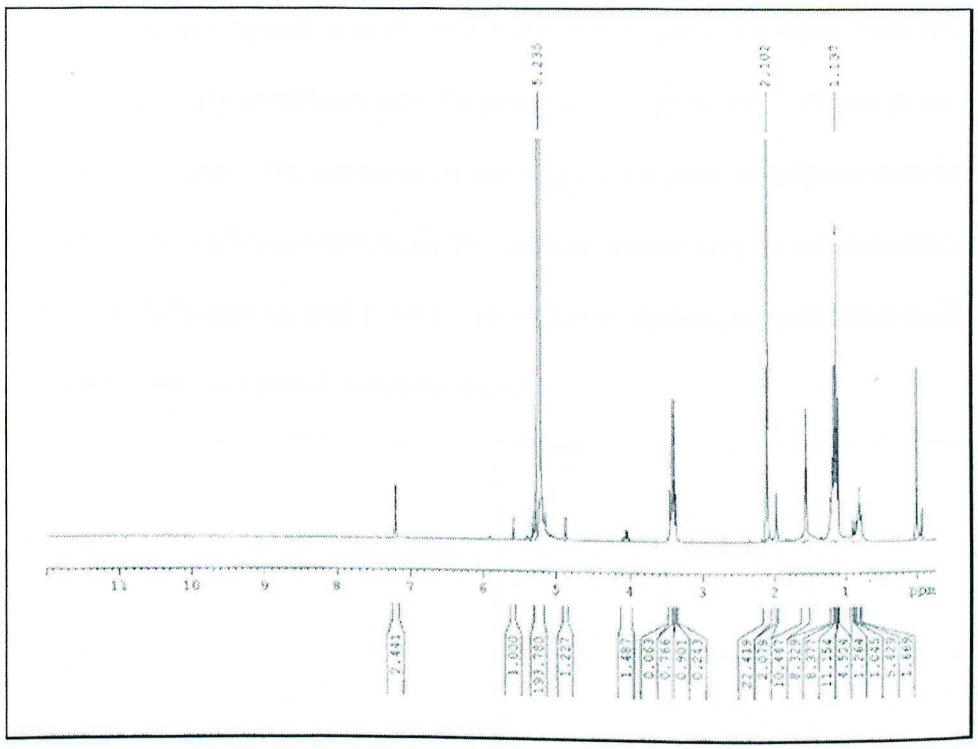

Figure 22. H NMR spectrum for Substrate B, sample 1. 1H NMR (300MHz, $CDCl_3$) δ (ppm): 5.60 (s, 1H), 4.89 (s, 1H), 3.42 (q, 2H), 1.98 (s, 2H), 1.57 (s, 10H), 1.15 (quin, 32H), 0.95 (d, 2H)

For substrate B sample 1, the highest frequency MW of a parent ion was 531.3 g/mol (Figure 21). This MW is approximately 60 g/mol less than the expected MW for Substrate B. This MW indicates that substrate B sample 1 is the diacylated form of Substrate B, as the removal of two methylene carbon ether groups gives an expected MW of 530.63 g/mol (2C + 4H + 2O). The fragment MW, 453.1 g/mol, is likely the fluorescein molecule with only one acylated attachment to it. Fluorescein has a MW of 332.31 g/mol and the acylated attachment (no methylene carbon ether) has a MW of 116.16 g/mol. When these are added, the sum is 448.47 g/mol; this total is close to 453.1 g/mol of the fragment.

The H NMR spectrum in Figure 22 expectedly shows proton peaks mostly on the lower, right side. The peak representing the proton at the tertiary carbon between the

ethyl groups can be distinguished as the quadruplet 2H peak at 3.42ppm. The other alkane protons are likely combined into the peaks at 1.57ppm and 1.15ppm in an indistinguishable pattern. The presence of the singlet, 1H peak at 5.6ppm does not likely indicate a methylene carbon proton since this feature would only be indicated by a 2H peak. Both the LCMS spectra and H NMR spectrum complement each other to show that the first sample is the diacylated substrate product.

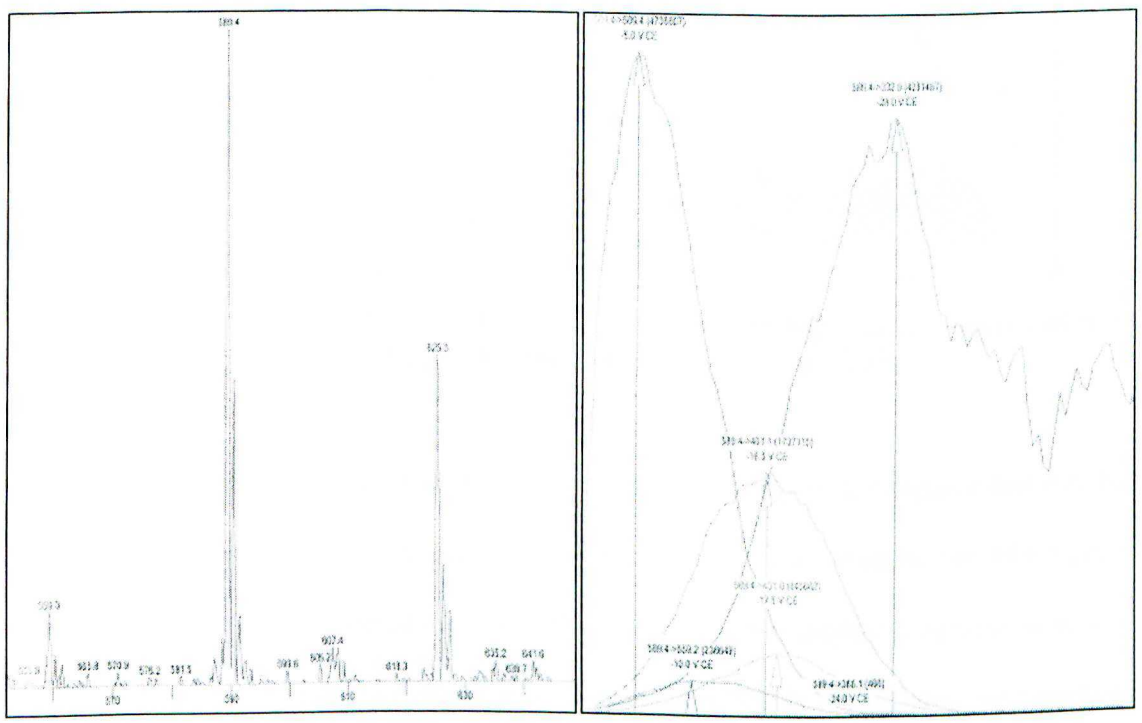

Figure 23. LCMS spectra of Substrate B, sample 2. The MW for parent ions (left) and the MW frequency of daughter fragments (right). The most frequent parent ion MW is 589.4 g/mol and the top three fragment MWs are 332.1 g/mol, 451.1 g/mol, and 431 g/mol.

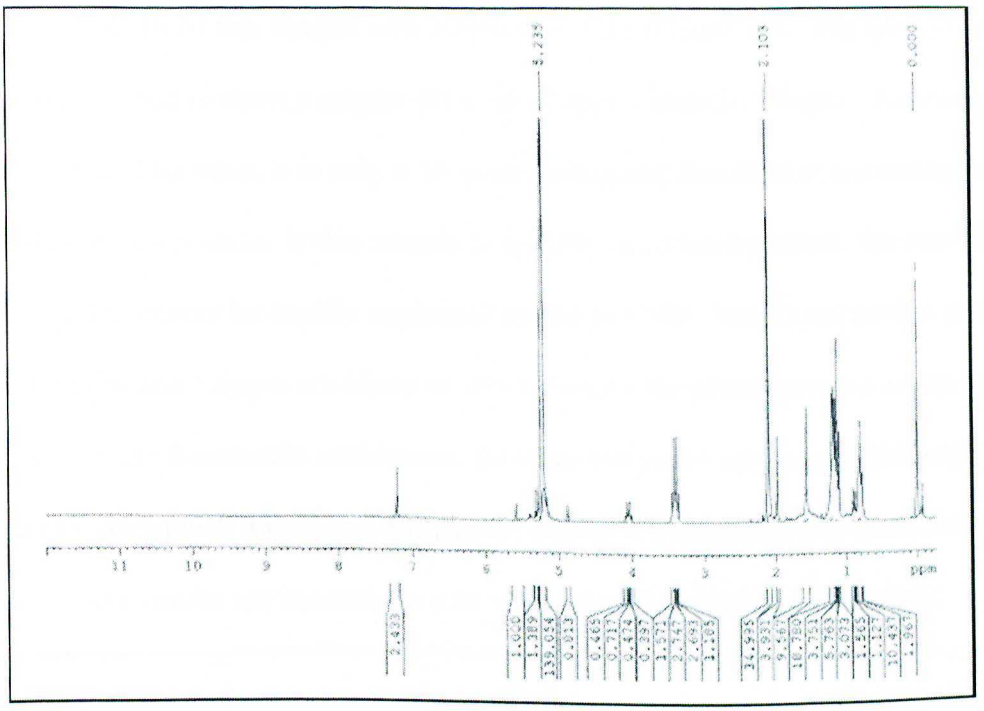

Figure 24. H NMR spectrum for Substrate B, sample 2. 1H NMR (300MHz, $CDCl_3$) δ (ppm): 5.60 (s, 1H), 4.89 (s, 1H), 3.42 (q, 8H), 1.98 (s, 4H), 1.60 (s, 10H), 1.19 (quin, 32H), 0.90 (d, 2H)

The LCMS spectra from sample 2 of Substrate B (Figure 23) suggest that it is the dialkylated substrate. The parent ion MW most detected in this solution was 589.4 g/mol and is very close to the expected MW of 590.63 g/mol. The daughter fragment MW of 332.9 g/mol represents the fluorescein molecule by itself (MW 332.31) and the fragment MW of 451.1 g/mol likely represents the dialkylated substrate with one alkylated attachment removed. When the MW of the alkylated attachment is subtracted from the parent ion MW (589.4 g/mol - 129.16 g/mol), the total difference of 460.2 g/mol is somewhat close to the observed fragment MW. The other fragment MW, 431 g/mol, is not easily explained because a leaving group of about 20 g/mol is not readily seen on the monoalkylated substrate of MW 451.1 g/mol.

The H NMR of this sample tells a different story (Figure 24). The dialkylated substrate is expected to show a singlet 4H peak at approximately 5.8ppm. Although a peak exists in this location, it is only a 1H peak, indicating that there is not methylene carbon ether in this product. If this sample is actually diacylated product, the extra MW shown in the LC cannot be readily explained by the H NMR. The alkane proton peaks between 0.9ppm and 1.6ppm are likely to still represent the alkane protons on the two ethyl groups in the fluorescein attachment. Because the peaks are noisy in this region, it is difficult to distinguish any major differences between these peaks and those for the first substrate. These results are inconclusive as to the identity of the second sample.

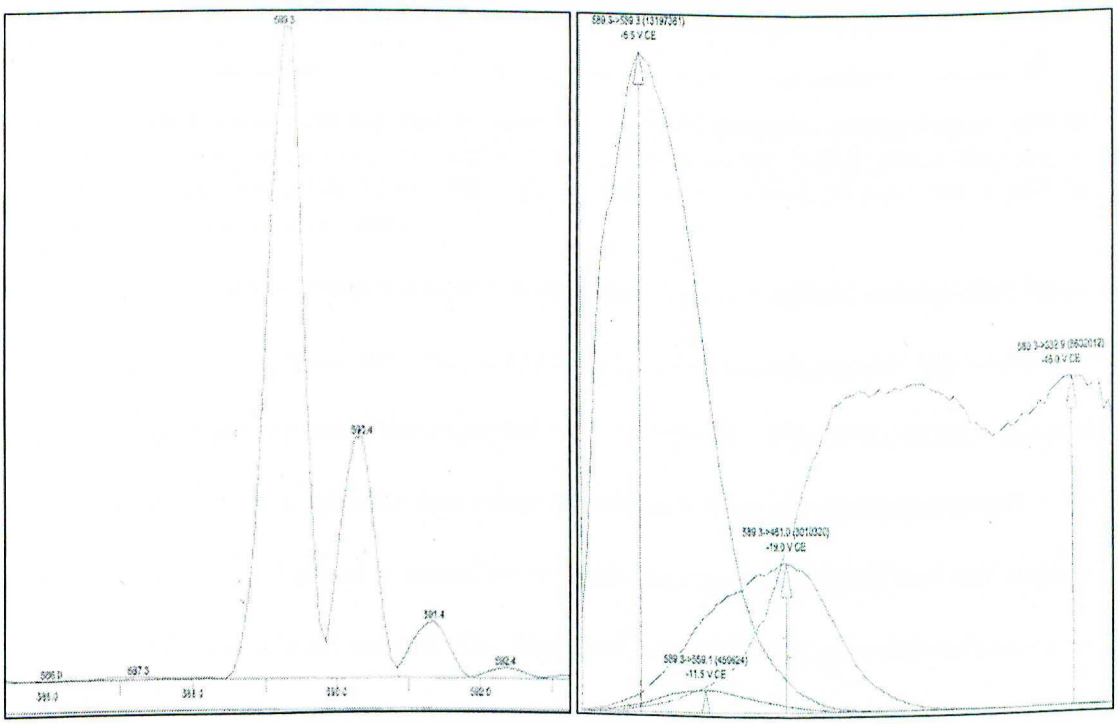

Figure 25. LCMS spectra of Substrate B, sample 3. The MW for parent ions (left) and the MW frequency of daughter fragments (right). The most frequent parent ion MW is 589.3 g/mol and the top three fragment MWs are 332.9 g/mol, 461.1 g/mol, and 559.1 g/mol.

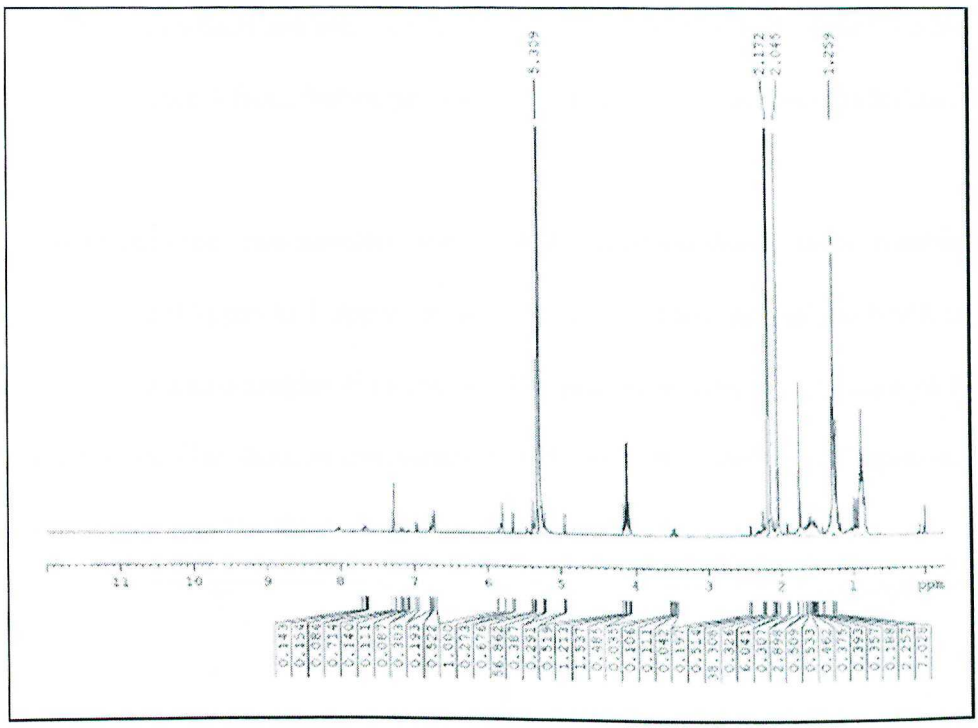

Figure 26. H NMR spectrum for Substrate B, sample 3. 1H NMR (300MHz, $CDCl_3$) δ (ppm): 8.01 (d, 2H), 7.66 (ddd, 1H), 7.16 (d, 1H), 6.97 (s, 1H), 6.81-6.71 (m, 2H), 5.81 (s, 4H), 5.66 (s, 1H), 5.39 (s, 2H), 5.23 (s, 1H), 4.95 (s, 1H), 3.48 (q, 1H), 2.25 (s, 2H), 2.01 (s, 18H), 1.75 (s, 12H), 1.58 (sex, 12H), 1.35-1.16 (m, 38H), 0.97 (d, 5H), 0.93-0.80 (m, 24H).

The third sample from Substrate B provided results that best exemplified those of the desired substrate (Figure 25). As was the case in the second sample, the overall MW of 589.3 g/mol was very close the expected MW of 590.63. The MWs of the fragments were representative of fragments that could be released from the desired substrate. The fragment of MW 332.9 g/mol is representative of fluorescein by itself, and this fragment also appeared in the second sample. The fragment with MW 461 g/mol is indicative of the fluorescein with one alkylated attachment. If the MW of the carboxylic acid (116.16 g/mol) and the MW of the methylene carbon (14 g/mol) are subtracted from the MW of the expected substrate (590.63 g/mol), the difference is 460.47 g/mol; this is very close to the fragment's MW of 461 g/mol. The final fragment of MW 559.1 g/mol is approximately 30 g/mol less than the expected MW of Substrate B; this likely indicates a

product fragment in which one ethyl group (29 g/mol) was removed. Based on the LCMS results, sample 3 from Substrate B was identified as the desired dialkylated product.

As with the other two samples, the H NMR spectrum shows noisy, combined proton peaks in the 0.8ppm to 1.9ppm range (Figure 26). This sample's H NMR is different because it had a singlet 4H peak at 5.81ppm, indicating the presence of two methylene carbons. This data, in combination with the MWs from the LC spectra, provide confidence that the third sample is dialkylated Substrate B.

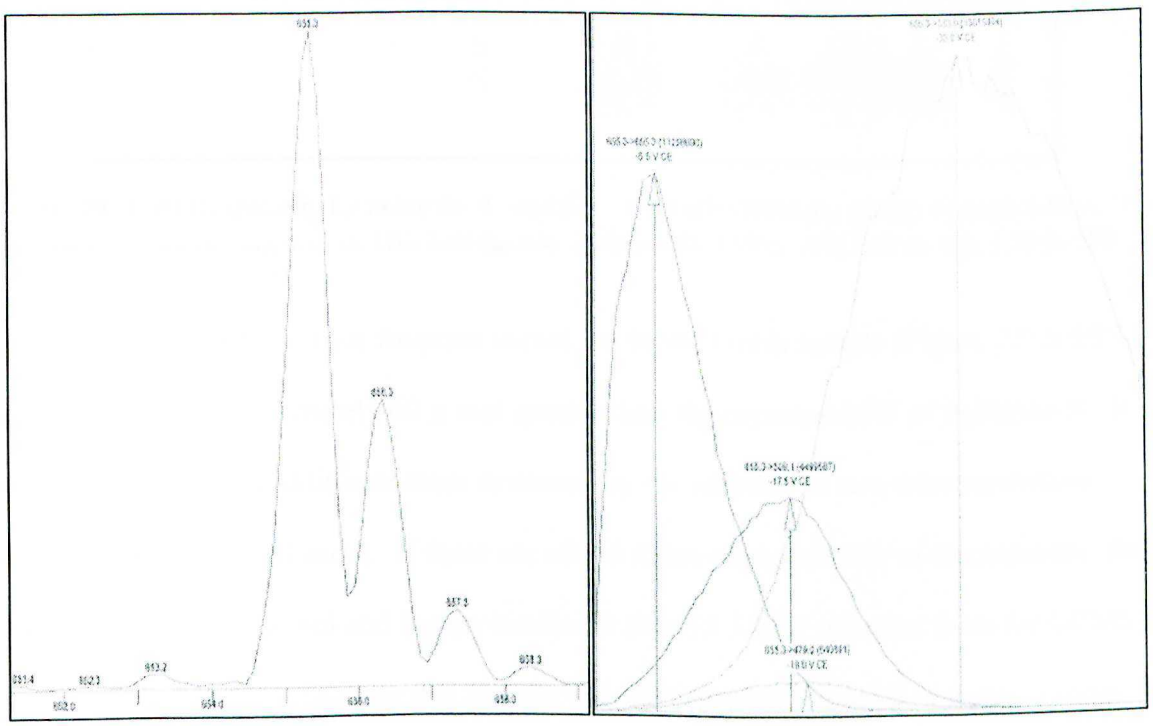

Figure 27. LCMS spectra of Substrate B, sample 4. The MW for parent ions (left) and the MW frequency of daughter fragments (right). The most frequent parent ion MW is 655.3 g/mol and the top three fragment MWs are 333 g/mol, 555.3 g/mol, and 509.1 g/mol.

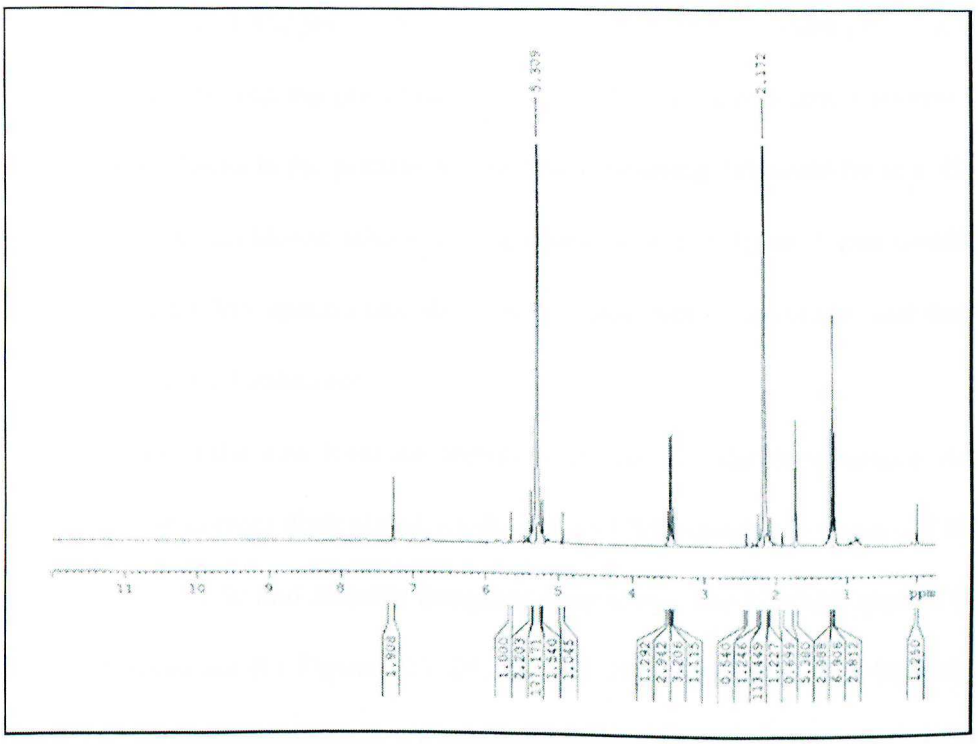

Figure 28. H NMR spectrum for Substrate B, sample 4. 1H NMR (300MHz, $CDCl_3$) δ (ppm): 5.66 (s, 1H), 5.39 (s, 2H), 5.23 (s, 1H), 4.95 (s, 1H), 3.48 (q, 8H), 2.25 (s, 1H), 2.09 (s, 1H), 1.92 (s, 1H), 1.73 (s, 5H), 1.58 (sex, 12H), 1.2 (t,12H).

The MW of the most frequent parent ion in the fourth sample (Figure 27) is 655.3 g/mol. This is approximately 60 g/mol greater than the expected MW of Substrate B. It is proposed that this additional mass is caused by the addition of two extra methylene carbon ethers (30 g/ml each). If these are added to the expected MW of the substrate, the total MW is 650.63 g/mol and is very similar to the 655.3 MW obtained from the LCMS spectra. This is logical because this sample was the last to be removed from the column during separation and was therefore the more polar product obtained; additional ethers increase the overall polarity of the molecule, and this is assumed to be the cause of this observed polarity increase. Clearly, this is not the desired substrate due to this high MW.

The H NMR spectra in Figure 28 also support the determination that the product in this sample contains additional methylene carbon ethers. The peaks at 5.66ppm,

5.39ppm, 5.23ppm, and 4.95ppm could all indicate the presence of methylene carbon ether. These total to 5H, and the presence of four methylene carbon ethers should be represented by 8H. There is the possibility that the remaining 3H could be in a fifth peak that is obscured by the dichloromethane contaminant peak at 5.4ppm. Upon combining the results from the LCMS spectra and the H NMR spectrum, it is evident that this product is not the desired substrate.

Based on all of the data from the Substrate B samples, the third sample was determined to be the correct dialkylated product. Its LCMS spectra in Figure 25 clearly showed the expected MW and feasible daughter fragments. The H NMR spectra from the four samples of Substrate B (Figures 22, 24, 26, and 28) exhibited similar behavior with key differences at the location where a methylene carbon ether would be located. As with the H NMR analysis for the Substrate A samples, desired dialkylated products are expected to show a singlet, 4H peak at approximately 5.8ppm. The only spectra that fits these criteria is for Substrate A sample 3. This nicely complements the data from the LCMS spectra analysis. Although the true identities of the second and fourth samples are unknown, an assertion can still be made that the third sample is the dialkylated Substrate B. Since this is the only sample between the two sets of synthetic products that has been identified as one of the substrates of interest, it is the only one that can be confidently used in enzymatic kinetics experiments.

Enzymatic Kinetics

Upon identification of substrate B by LCMS and H NMR spectroscopy, three functional assays were performed. These assays exposed the substrate to one of two enzymes in the lab. During these tests, the esterase would cleave the substrate at its ester bonds, forming two carboxylic acids and a fluorescein molecule in its fluorescent state. This is shown in Figure 29 in which substrate B interacts with Rv0045c to produce a fluorescein molecule and 2-ethylbutyric acid. The increase in fluorescence upon ester hydrolysis of Substrate B was monitored over time. Each enzyme assay utilized a 96 well microplate to test eight substrate concentrations ranging from 0.05μM to 100μM in 1:2 dilutions. Initial velocities for fluorophore hydrolysis were measured in triplicate to determine the error for the measurement. As fluorescence is proportional to the concentration of cleaved substrate, a fluorescein standard curve was used for each assay to convert fluorescence change over time into concentration change over time.

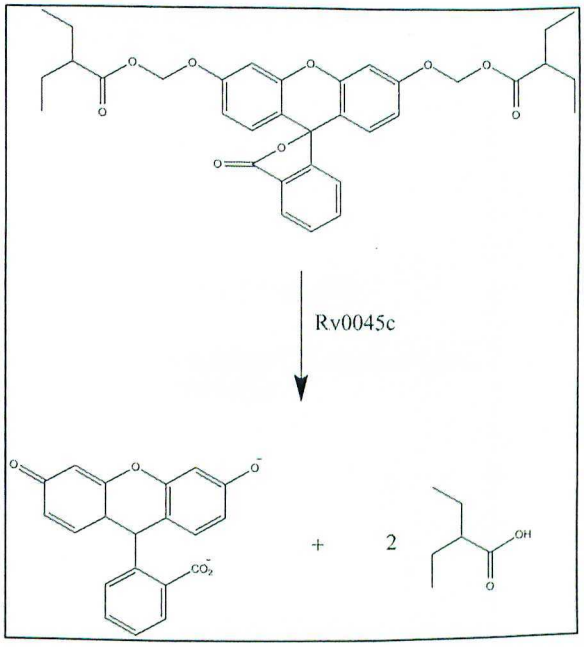

Figure 29. Hydrolysis of esters within fluorogenic substrate. Schematic of the cleavage at the ester bonds in substrate B caused by RV0045c. Products include fluorescein (in its fluorescent state) and 2-ethylbutyric acid.

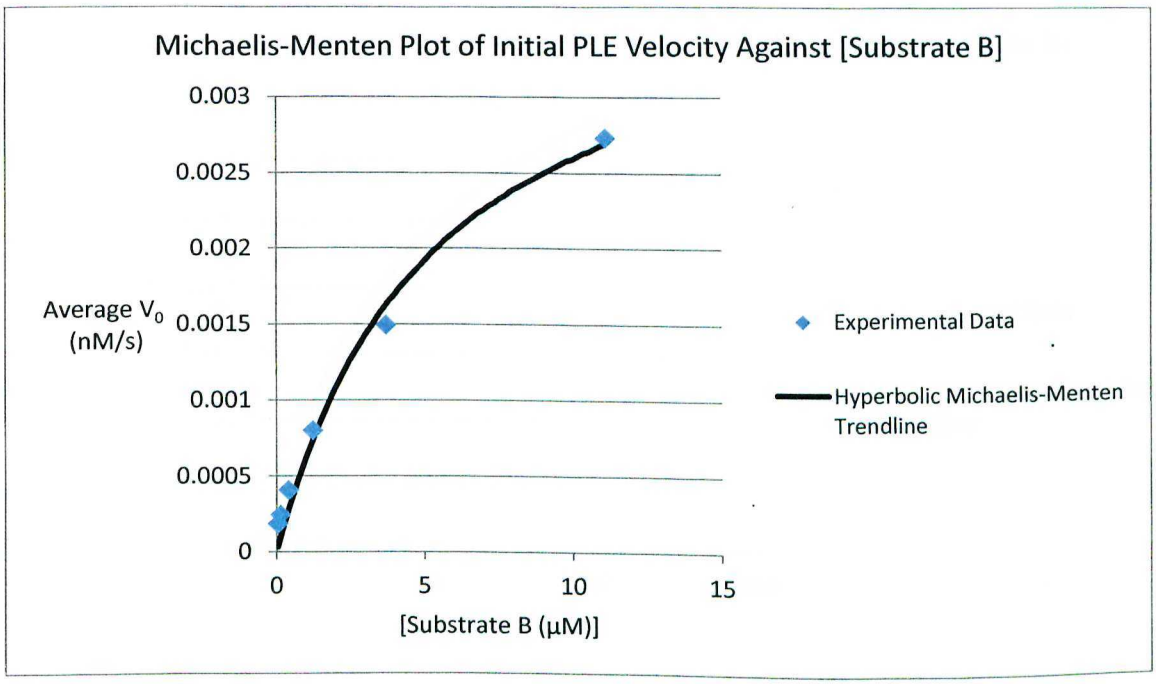

Figure 30. Michaelis-Menten plot of Initial PLE velocity versus [Substrate B]. Data from substrate B hydrolysis by pig liver esterase (PLE) is shown in blue dots, and the trendline generated in the Origin 6.1 program through a fit against the hyperbolic Michaelis-Menten equation is shown in black. The generation of this trendline provided best fit values for V_{max} and K_m.

To show that Substrate B behaved like a latent fluorophore, the background hydrolysis rate and enzymatic hydrolysis rates were compared. The rate of background substrate hydrolysis was minimal and Substrate B is highly stable in solution (data not shown). This stability in solution is related to the presence of the methyl ether with its increased pK_a. In the presence of pig liver esterase (PLE), a common, general esterase, Substrate B showed a rapid increase in fluorescence that scaled with concentration and the results could be fitted to the Michaelis-Menten equation (Figure 30; Table 1). Together these results confirm that Substrate B has the properties of an alkylated latent fluorophore with low background hydrolysis but rapid hydrolysis in the presence of an esterase.

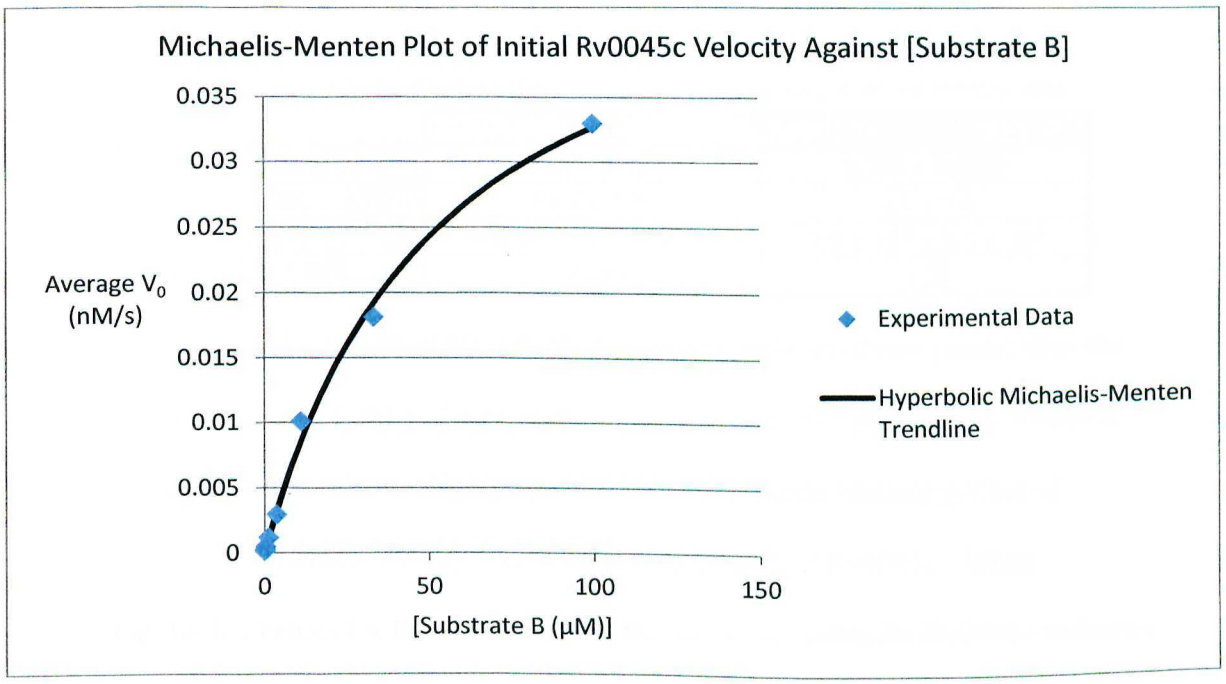

Figure 31. Michaelis-Menten plot of Initial Rv0045c velocity versus [Substrate B]. Data from substrate B hydrolysis by the enzyme of interest, Rv0045c, is shown in blue dots, and the trendline generated in the Origin 6.1 program through a fit against the hyperbolic Michaelis-Menten equation is shown in black. The generation of this trendline provided best fit values for V_{max} and K_m.

Now that the chemical identity and latent fluorophore properties of Substrate B were confirmed, the substrate was tested for activity against Rv0045c (Figure 31). Although the initial velocity for the hydrolysis of Substrate B by Rv0045c was not as high as for the hydrolysis by PLE, significant hydrolysis was still observed. To get a quantitative grasp on the differences between the enzymatic behavior of Rv0045c and PLE enzymes, Origin 6.1 computer software was used to fit the data from the assays to the Michaelis-Menten equation (Equation 1, background). This data fit calculated the V_{max} and K_m for each of the tested enzymes; k_{cat} and specificity constant values were determined by using these obtained values with the Michaelis-Menten equations provided in the Background portion of this thesis (Equation 1, Equation 2). All Michaelis-Menten kinetic values are shown in Table 1.

Table 1. Michaelis-Menten kinetic values based upon curve fitting of initial velocity data

	PLE	Rv0045c
V_{max} (nM/s)	0.05 ± 0.003	0.004 ± 0.0005
K_m (μM)	53.5 ± 7.06	5.5 ± 1.6
k_{cat} (s^{-1})	0.001 ± 6.9 x 10^{-5}	1.6 x 10^{-5} ± 2.1 x 10^{-6}
k_{cat}/K_m (s^{-1} M^{-1})	20.99	2.95

PLE has a V_{max} of 0.05 nM/s, which is approximately ten times greater than the V_{max} of Rv0045c at 0.004 nM/s. These relatively lower rates of hydrolysis by Rv0045c suggests that Substrate B is not highly specific for the substrate binding pocket of Rv0045c and only interacts weakly with the binding pocket of Rv0045c. When comparing the K_m values for the two enzymes, the lower K_m value for Rv0045c indicates that the substrate binds better to Rv0045c than it does to PLE. This lower K_m value indicates that a lower concentration of substrate is needed to obtain half of the enzyme's maximum velocity. Although this is the case, the k_{cat} value, which represents the turnover rate, is much lower for Rv0045c than PLE. The determining factor of which of the two enzymes was more specific to substrate B was the specificity constant (k_{cat}/K_m) since it weighed the k_{cat} and K_m values against one another. Because the k_{cat} value of Rv0045c was so much less than that of PLE, the specificity constant was higher for PLE than for Rv0045c. Since PLE is a common enzyme and substrate B has a lower specificity constant for Rv0045c than for PLE, it can be concluded that the substrate is currently far from specific for this TB lipase; it may, however, represent a potential starting point for future prodrug developments. The characterization of Rv0045c should also be repeated before making any firm conclusions about its specificity to the substrate.

Conclusion

In order to examine whether Rv0045c would show specificity toward either steric or flexible substrates, two latent fluorophores were created with unique acetoxymethyl ether attachments. Upon completion of the synthetic steps necessary for their creation, only one desired flexible substrate, fluorescein di(ethylbutyloxymethyl ether), was successfully identified *via* liquid chromatography mass spectrometry and NMR spectroscopy. Enzymatic and background hydrolysis tests showed that the substrate displayed very little specificity to the Rv0045c enzyme. Comparisons between steric and flexible substrate specificity cannot be made until a successful synthesis of the desired steric substrate, di(phenylbutyloxymethyl ether), or some other steric substrate can be accomplished.

Future Directions

To accomplish the initial purpose of this thesis research, it would be necessary to perform the synthesis of the steric substrate once more. Upon completion of the repeated synthesis of this substrate, the column chromatography would be conducted using a more nonpolar solvent to slow the movement of synthesized products through the column and increase the separation between them. This work would additionally benefit from further research into the specificity of the flexible substrate that was successfully synthesized. This substrate should be tested on other bacterial esterases from the causative agents of cholera, tularemia, and others in order to gain a better understanding of its relative specificity to Rv0045c in *M. tuberculosis*.

Materials and Methods

Materials

All chemical reagents used for the syntheses and experiments in this thesis were purchased from Sigma Aldrich and chosen based on the highest grade necessary for the purposes of this research. Pig liver esterase (PLE) used in substrate testing was from Sigma, and Rv0045c (esterase from *Mycobacterium tuberculosis*) was purified using the previously published procedure (Guo *et al.*, 2010). Phosphate buffered saline buffer (PBS) pH 7.4 was made using 8g NaCl, 1.44g Na_2HPO_4, 0.27g KH_2PO_4, 0.2g KCl.

Selection of Substrates

The synthesis of two fluorogenic enzyme substrates were attempted using identical methods; these two substrates differed only in the carboxylic acid that was chosen for the initial synthesis step. The two carboxylic acids chosen were 2-phenylbutyric acid and 2-ethylbutyric acid and the major expected difference between the two products was their inherent flexibility (Figure 31 (a) and (d)). 2-ethylbutyric acid is likely highly flexible due to its two unconstrained ethyl chains off of a tertiary carbon; 2-phenylbutyric acid contains one ethyl chain and one benzene ring connected to a tertiary carbon, thus causing the acid to be bulky and comparatively more sterically hindered than 2-ethylbutyric acid. Although bulky, each of these carboxylic acids is fairly short in comparison to standard long lipid chains, making them good candidates as substrates for the esterases used in this research. The following procedure details the general synthetic methods used to synthesis both fluorogenic enzyme substrates (fluorescein di(phenylbutyloxymethyl) ether and fluorescein di(ethylbutyloxymethyl ether), Figure 31

(c) and (f), respectively) using the term 'carboxylic acid' in place of 2-phenylbutyric acid and 2-ethylbutyric acid.

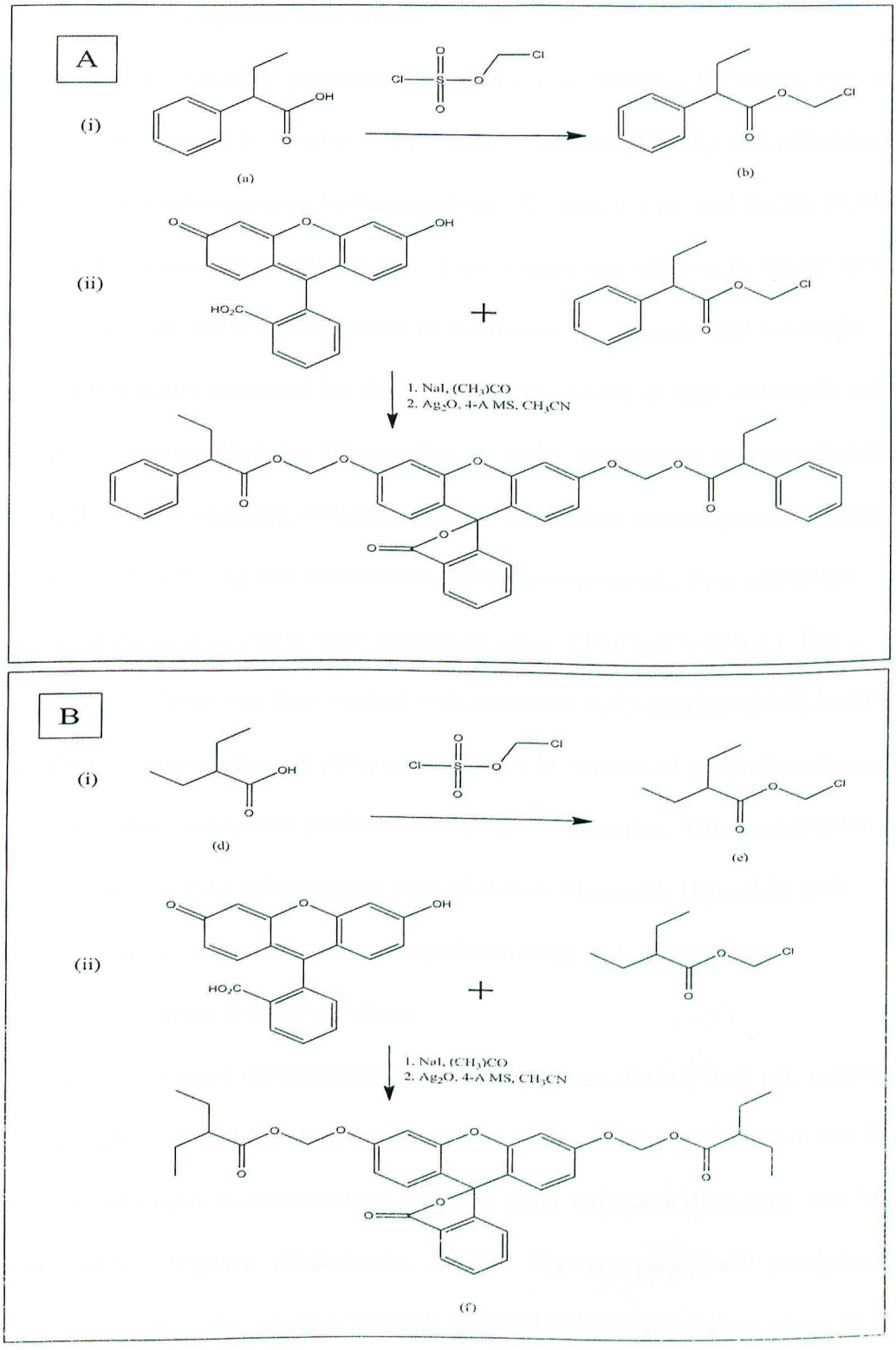

Figure 32. Fluorogenic Substrate Synthesis of (A) fluorescein di(phenylbutyloxymethyl ether) and (B) fluorescein di(ethylbutyloxymethyl ether). Each synthesis has two main steps, (i) methyl chlorination and (ii) fluorescein alkylation.

Synthesis of chloromethylated acid products (Figure 31, i)

Initially, a solution of chloromethyl chlorosulfate (600mg, 2.43mmol, 1eq) and carboxylic acid (a) or (d) (1.5 eq) in 5 mL dichloromethane (CH_2Cl_2) was added to a solution of tetrabutylammonium hydrogensulfate (82.5mg, 0.1 eq) and K_2CO_3 (1.345g, 4 eq) in 10mL DI water and 10 mL CH_2Cl_2. This mixture was allowed to stir for 12 hours at room temperature and was then placed in the freezer for storage. This overnight synthesis theoretically produced the choromethylated version of each carboxylic acid: 2-phenylbutyric chloromethylester (Figure 31A, b) and 2-ethylbutyric chloromethylester (Figure 31B, e). This resulting colloidal mixture was poured into a separatory funnel and the organic layer (CH_2Cl_2) was immediately removed and saved. Two additional extractions of the aqueous layer were performed using CH_2Cl_2 (2 x 15 mL). The combined organic layer was then washed with saturated NaCl solution (30mL) and dried using $MgSO_4$. Filtration through silica gel was used to remove all particulates from the solution, and condensation was performed by rotary evaporation. This concentrated product was the resulting chloromethyl ester shown in Figure 31, (b) and (e) and concluded the initial step of the substrate synthesis (step (i) in Figure 31).

Synthesis of fluorogenic enzyme substrate

The concentrated chloromethyl ester product was diluted in 5 mL acetone and mixed overnight with NaI (361mg, 2.41mmol) as shown in Figure 31 A(ii)-1 and B(ii)-1; this reaction was used to substitute the chlorine atom with an iodine atom and form an insoluble NaCl precipitate (Finkelstein, 1910). The resulting NaCl precipitate was removed by filtering the product through cotton loaded into a glass pipet in which pressure was occasionally applied. Excess acetone was applied to the solution as a wash

in order to ensure all product was collected, and the filtered iodomethyl ether product was concentrated used the rotovap. The concentrated iodinated ester was taken up in 2 x 2mL additions of acetonitrile. After being transferred to a scintillation tube, fluorescein (200mg, 0.602 mmol), 4-A molecular sieves (200mg) and anhydrous Ag_2O (348mg, 1.5 mmol) were added; this mixture was stirred for 7 days at room temperature to complete the substrate synthesis. The substrate solution was diluted with 100mL CH_2Cl_2 and vacuum filtered through a layer of Celite to remove all particulates. The filtered product was concentrated using the rotovap.

Column chromatography

To separate the correctly dialkylated fluorogenic substrate from other reaction products (diacylated and monoacylated-monoalkylated), silica gel column chromatography was performed. TLC plates were used to investigate the relative polarity of each potential product and determine an appropriate solution ratio of hexane, ethyl acetate, and dichloromethane to separate them *via* column chromatography. An RF value between 15 and 20% was desired such that products with slight differences in polarity could be separated from one another; this precaution prevented all of the products from coming off of the column together in a highly polar solution. Five solutions (10mL each) of differing hexane:ethyl acetate:dichloromethane ratios were created and used on substrate-dotted TLC strips as follows: 8:0:2, 7:1:2, 6:2:2, 5:3:2, and 5:1:4. The least polar solution (8:0:2) showed almost no band movement, and the slightly more polor solution (7:1:2) provided RF values around 20% for both substrates. The other three solutions had RF values between 50-90% and were thus eliminated from consideration as a column chromatography solution. Celite (3g) and substrate were put into CH_2Cl_2

solution. Celite, being a porous powder, absorbed the substrate solution, and took up the product before CH_2Cl_2 was removed by vacuum in the rotovap. At this point, the Celite not only acted as a convenient way to transfer the substrate as a powder (rather than oily concentrated solution), but it also acted as a solid particulate filter within the column. Columns were prepared using an amount of silica gel equal to one-tenth of the substrate's molecular weight (c: 68.67g, f: 59.06g) and enough sand to create a 2 inch layer above the substrate powder. The initial column solution was composed of 8:0:2 solution to keep from early flow through the column. After 300 mL of this solution was run through the column, the solution was switched to a 7:0.5:1.5 ratio; this made the solution slightly more polar to allow the potential products to move down the column. Product bands could be seen using a UV light that detected the fluorophore substrate component; once clear product band separation was seen using this method, polarity was again increased slightly to a 7:1:2 solution in order to speed product movement through the column. Once the UV light detected product bands within an inch of the bottom of the column, samples were collected until all bands had passed through. Synthesis A (Figure 1A) produced four product bands off of the column and synthesis B (Figure 1B) produced three product bands. Each substrate purification produced about 60, 5 mL fractions.

Chemical characterization of final synthetic products

The fractions collected from the column chromatography were initially analyzed by TLC. Every third fraction was tested in a 7:0.5:1.5 TLC solution in order to determine which of the fractions contained products. Because the solution was slightly polar, fractions containing more than one product showed two separate product bands on the plate. UV light allowed for these bands to be seen, and a heat gun was used to find bands

with denatured product (heat would cause these bands to turn pale yellow). After performing initial TLC testing on one third of the fractions, the untested fractions surrounding tested, product-containing fractions were then run on a TLC strip; this prevented the testing of all sixty fractions if it was unnecessary. Three types of products were expected with slightly different polarities: di-alkylated fluorescein (least polar), mono-alkylated mono-acylated fluorescein, and di-acylated fluorescein (most polar). The desired product, di-alkylated fluorescein, was expected to be one of the last, least-polar detected bands in the collected fractions. Fractions containing identical, singular bands on the TLC strips were pooled together and then characterized by H NMR and LCMS with substrates put into deuterated chloroform and acetonitrile, respectively, to determine which band was the desired di-alkylated substrate. The LCMS analysis was performed on a triple quadripole. Solution was injected to the machine at a steady rate using a syringe. Liquid chromatography was used initially within the machine to separate molecules within the solution by polarity. Mass spectrometry was performed within the triple quadripole to determine the MW of the parent ion; the MW is based on the mass to charge ration (m/z) that is determined by the molecule's travel time. Finally, another mass spectrometry was performed in which the parent ion was split into fractions by argon; this allowed for the determination of daughter ion MW. Test results from H NMR and LCMS analysis are shown in detail in the Results and Discussion section.

Fluorescence testing against PLE and Rv0045c

Using the mass of each substrate of interest, dimethylsulfoxide solvent was added to each to create either a 10mM or 100mM solution. Using a fluorometer, PBS solution was used as a control solution in emission and excitation scans from 300 to 600 nm. The

substrate solutions were subjected to a 1:200,000 dilution in PBS, and emission and excitation scans were performed with the emission maximum set at 520nm and the excitation maximum set at 494nm. Pig liver esterase (PLE) was added in 10uL increments until increased activity levels were observed. Additional scans were performed on PLE diluted in PBS solution to ensure that the increased activity was substrate-dependent.

Kinetic testing against PLE and Rv0045c

Three functional assays were performed in which fluorescence due to ester hydrolysis was monitored over time. A 96 well microplate was utilized to create and test eight substrate concentrations ranging from 0.05μM to 100μM in 1:2 dilutions in PBS (150 μg/mL) + BSA (0.1 mg/mL) at pH 7. To the substrate dilutions, an esterase (5μL of 150 μg/mL) was added for the PLE and Rv0045c functional assays, and PBS buffer was added for the background substrate hydrolysis assay. Immediately after the addition of enzyme, the microplate was placed into the fluorescent plate reader. The fluorescent plate reader was set to measure fluorescence every 15 seconds for four minutes at an excitation wavelength of 485nM and an emission wavelength of 520nm. A fluorescence standard solution was prepared for each assay by creating eight known fluorescein solutions with concentrations ranging from 2.3nM to 300 nM in 1:1 dilution in PBS buffer. From this data, a fluorescein standard curve was made, which converted fluorescence into the concentration of cleaved substrate. Initial velocity (V_0) values were calculated by determining the slope of the cleaved substrate concentration against time for each substrate concentration. Initial velocities were averaged over the triplicate substrate measurements to account for error. Using the Origin software program, the

initial velocities and substrate concentrations were fitted to a Michaelis-Menten curve, and the V_{max} and K_m values and error measurements were recorded.

References

Cole, S. T., R. Brosch, J. Parkhill, T. Garnier, and C. Churcher. "Deciphering the Biology of Myscobacterium Tuberculosis from the Complete Genome Sequence." *Nature* 393 (1998): 537-44. Print.

Cotes, K., N'Goma, J.C., Dhouib, R., Douchet, I., Maurin, D., Carriere, F., Canaan, S. (2008) Lipolytic enzymes in *Mycobacterium tuberculosis. Appl Microbiol Biotechnol* 78:741-749.

Finkelstein, Ber. Dtsch. Chem. Ges., 1910, 43, 1528.

"Fluorescence." *Encyclopædia Britannica. Encyclopædia Britannica Online Academic Edition*. Encyclopædia Britannica Inc., 2012. Web. 16 Mar. 2012. <http://www.britannica.com/EBchecked/topic/211338/fluorescence>.

Gottlieb, Hugo E., Vadim Kotlyar, and Abraham Nudelman. "NMR Chemical Shifts of Common Laboratory Solvents as Trace Impurities." *The Journal of Organic Chemistry* 62.21 (1997): 7512-515.

Guo, J, Zheng X, Xu L, Liu Z, Xu K, et al. (2010) Characterization of a Novel Esterase Rv0045c from *Mycobacterium tuberculosis*. PLoS ONE 5(10): e13143. doi:10.1371/journal.pone.0013143

Holmquist, M.2000. Alpha/Beta-Hydrolase Fold Enzymes: Structures, Functions and Mechanisms. *Current Protein and Peptide Science*. 1:209-235.

Hughes, Andrew B. *Amino Acids, Peptides, and Proteins in Organic Chemistry*. Weinheim: Wiley-VCH, 2009.

Lavis, Luke D., and Ronald T. Raines. "Bright Ideas for Chemical Biology." *ACS Chemical Biology* 3.3 (2008): 142-55.

Lavis, Luke D., Tzu-Yuan Chao, and Ronald T. Raines. "Synthesis and Utility of Fluorogenic Acetoxymethyl Ethers." *Chemical Science* (2010). Print

Menten, L., Michaelis, M.I. (1913). Die Kinetik der Invertinwirkung. Bioche Z 49:333–369

Nelson, David L., Michael M. Cox, and Albert L. Lehninger. *Principles of Biochemistry*. New York, NY: Freeman, 2008. Print.

Power, Nicholas P., Donald Bethell, Lee Proctor, Elliot Latham, and Paul Dawson. "Chloromethyl Chlorosulfate: A New, Catalytic Method of Preparation and Reactions with Some NucleophilesElectronic Supplementary Information (ESI) Available: Tables S1?S4 and Scheme S1. See Http://www.rsc.org/suppdata/ob/b4/b403338h/." *Organic & Biomolecular Chemistry* 2.10 (2004): 1554. Print.

Rautio, Jarkko, Hanna Kumpulainen, Tycho Heimbach, Reza Oliyai, Dooman Oh, Tomi Järvinen, and Jouko Savolainen. "Prodrugs: Design and Clinical Applications." *Nature Reviews Drug Discovery* 7.3 (2008): 255-70. Print.

Schicher, Maximilian, Maria Morak, Ruth Birner-Gruenberger, Heidemarie Kayer, Bojana Stojcic, Gerald Rechberger, Manfred Kollroser, and Albin Hermetter. "Functional Proteomic Analysis of Lipases and Esterases in Cultured Human Adipocytes." *Journal of Proteome Research* 9.12 (2010): 6334-344. Print.

Singh, G., Singh, G., Jadeja, D., Kaur, J. 2010. Lipid hydrolyzing enzymes in virulence: *Mycobacterium tuberculosis* as a model system. *Critical Reviews in Microbiology* 36(3): 259-269

Valeur, Bernard, and Mario N. Berberan-Santos. "A Brief History of Fluorescence and Phosphorescence before the Emergence of Quantum Theory." *Journal of Chemical Education* 88 (2011): 731-38.

West, Nicholas P., Katie M. Cergol, Millie Xue, Elizabeth J. Randall, Warwick J. Britton, and Richard J. Payne. "Inhibitors of an Essential Mycobacterial Cell Wall Lipase (Rv3802c) as Tuberculosis Drug Leads." *Chemical Communications* 47 (2011): 5166-168.

World Health Organization. (2010) [accessed February 26 2011] *Tuberculosis Global Facts*. Available from http://www.who.int/tb/publications/2010/factsheet_tb_2010rev21Feb11.pdf

World Health Organization.(2010) [accessed February 26 2011] Tuberculosis. Available from http://www.who.int/topics/tuberculosis/en.

Zheng X, Guo J, Xu L, Li H, Zhang D, et al. (2011) Crystal Structure of a Novel Esterase Rv0045c from *Mycobacterium tuberculosis*. PLoS ONE 6(5): e20506. doi:10.1371/journal.pone.0020506

Acknowledgements

My utmost gratitude is extended to my professor and research advisor, Dr. Jeremy Johnson, without whom this thesis would not have been possible. I thank him for his patience and guidance throughout not only the research component of this work, but also throughout the duration of the writing process. In addition, I want to thank the Butler University Chemistry Department for both supporting this project and funding its presentation at the annual Indiana Academy of Sciences meeting. I would also like to acknowledge the Butler University Honors Department for their guidance and advice throughout the course of my undergraduate career. Finally, I wish to thank Dr. Geoffrey Hoops for his willingness to be my thesis reader.